

**BAŐKENT ÜNİVERSİTESİ**  
**FEN BİLİMLERİ ENSTİTÜSÜ**

**İNSAN SAĐLIĐI TAKİBİ İÇİN GİYİLEBİLİR SENSÖR**  
**VERİLERİNİN ÇOK KİPLİ VE FÜZYON TABANLI ANALİZİ**

**GÖKHAN MEMİŐ**

DOKTORA TEZİ  
2019



**İNSAN SAĞLIĞI TAKİBİ İÇİN GİYİLEBİLİR SENSÖR  
VERİLERİNİN ÇOK KIPLI VE FÜZYON TABANLI ANALİZİ**

**MULTIMODAL AND FUSION BASED ANALYSIS OF  
WEARABLE SENSOR DATA FOR HUMAN HEALTH  
MONITORING**

**GÖKHAN MEMİŞ**

Başkent Üniversitesi  
Lisansüstü Eğitim Öğretim ve Sınav Yönetmeliğinin  
BİLGİSAYAR Mühendisliği Anabilim Dalı İçin Öngördüğü  
DOKTORA TEZİ  
olarak hazırlanmıştır.

2019

“İnsan Saęlıęı Takibi İin Giyilebilir Sensör Verilerinin ok Kipli Ve Füzyon Tabanlı Analizi” başlıklı bu alıřma, jürimiz tarafından, 16/09/2019 tarihinde, **BİLGİSAYAR MÜHENDİSLİęİ ANABİLİM DALI 'nda DOKTORA TEZİ** olarak kabul edilmiřtir.

Başkan :  
( Prof. Dr. Adnan YAZICI )

Üye (Danıřman) :  
( Dr. Öğr. Üyesi Mustafa SERT )

Üye :  
( Prof. Dr. Hasan OęUL )

Üye :  
( Do. Dr. Murat KOYUNCU )

Üye :  
( Do. Dr. Kumru Didem ATALAY )

**ONAY**

..../..../.....

Prof. Dr. Ömer Faruk ELALDI  
Fen Bilimleri Enstitüsü Müdürü



**BAŞKENT ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ**  
**DOKTORA TEZ ÇALIŞMASI ORJİNALLİK RAPORU**

Tarih: 16.09.2019

Öğrencinin Adı, Soyadı: Gökhan MEMİŞ

Öğrencinin Numarası: 21310021

Anabilim Dalı: Bilgisayar Mühendisliği

Programı: Doktora

Danışmanın Unvanı/Adı, Soyadı: Dr. Öğr. Üyesi Mustafa SERT

Tez Başlığı: İnsan Sağlığı Takibi İçin Giyilebilir Sensör Verilerinin Çok Kipli ve Füzyon Tabanlı Analizi

Yukarıda başlığı belirtilen Doktora tez çalışmamın; Giriş, Ana Bölümler ve Sonuç Bölümünden oluşan, toplam 100 sayfalık kısmına ilişkin, 16/09/2019 tarihinde şahsım tarafından Turnitin adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı %9'dur.

Uygulanan filtrelemeler:

1. Kaynakça hariç
2. Alıntılar hariç
3. Beş (5) kelimedenden daha az örtüşme içeren metin kısımları hariç

“Başkent Üniversitesi Enstitüleri Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Usul ve Esaslarını” inceledim ve bu uygulama esaslarında belirtilen azami benzerlik oranlarına tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Öğrenci İmzası:

Onay

16.09.2019

Dr. Öğr. Üyesi Mustafa SERT

## TEŐEKKÜR

Yazar, bu alıőmanın gerekleőmesinde katkılarından dolayı, aőađıda adı geen kiői ve kuruluőlara itenlikle teőekkür eder.

Sayın Yrd. Do. Dr. Mustafa SERT'e (tez danıőmanı), tez aőamamda verdiđi destek, alıőmanın sonuca ulaőtırılması ve karőtılaőtılan gülüklerin aőtılmasında her zaman yardımcı ve yol gösterici olduđu, araőtırmacı olmam konusunda desteđi ve yönlendirmeleri, deđerli bilgilerini benimle paylaőtıđı, kendisine ne zaman danıőtısam bana kıymetli zamanını ayırıp sabırla bana faydalı olabilmek iin elinden geleni yaptıđı, her sorun yaőtadıđımda yanına ekinmeden gidebildiđim, samimiyetini esirgemeyen ve gelecekteki mesleki hayatımda da bana verdiđi deđerli bilgilerden faydalanacađımı düőtündüđüm iin,

Tez komitemin üyeleri olan Prof. Dr. Adnan YAZICI ve Prof. Dr. Hasan OĐUL'a tez alıőtımam esnasında destekleri ve yol gösterici oldukları iin,

Babam Mehmet MEMİŐ'e ve annem Hatice MEMİŐ'e tüm yaőtantım boyunca sınırsız sevgi, destek ve ilgi ile hep yanımda oldukları, eđitimi ve iyi insan olmayı ruhuma iőtledikleri iin,

Eőtım Saadet Aycan BÜKE MEMİŐ'e eőtım, dostum ve arkadaőtım olduđu iin, beni yalnız bırakmadıđı, beni desteklediđi, bana inandıđı, alıőtımam iin alan ve zaman yarattıđı iin,

Ođlum Gökay MEMİŐ'e ve kızım İlay MEMİŐ'e karőtı sahip olduđum evlat sevgim ile tez sürecindeki zorlukları daha rahat bir őekilde gememi sađladıkları iin,

## ÖZ

# İNSAN SAĞLIĞI TAKİBİ İÇİN GİYİLEBİLİR SENSÖR VERİLERİNİN ÇOK KIPLI VE FÜZYON TABANLI ANALİZİ

Gökhan MEMİŞ

Başkent Üniversitesi Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

Bu çalışmada, sağlık bilişimi alanında, yaşlı ve bakıma muhtaç kişilerin, sağlık durumlarının ve günlük aktivitelerinin takibi amacıyla, farklı sensörlerden toplanan verilerden anlamlı bilgilerin elde edilebilmesi için, çok kipli ve füzyon tabanlı analizlere dayalı akıllı yöntemler geliştirilmiştir. Giyilebilir sensör olarak nabız, EKG, oksijen doygunluğu kullanılmıştır. Bu sensörlerden toplanan veriler ile aritmi, uyku apnesi hastalıklarının takibi ve erken tanılmasının yanı sıra, cep telefonu içerisinde bulunan sensörlerden ivmeölçer, manyetik alan ölçer, jiroskop ve eğimölçer sensörleri ile izlenen kişinin fiziksel aktivitesi incelenmiştir. Tez çalışması kapsamında, 12 farklı kişiden açık hava ile kapalı havada bulunma durumuna göre fiziksel aktivite verisi toplanmıştır. Toplanan veriler ile kişinin fiziksel aktivitesini nerede yaptığı bulunmuştur. Bu aşamada standart sapma ve üstel hareketli ortalama tabanlı yeni öznelikler oluşturulmuştur. Kullanılan yeni öznelikler ile yapılan çalışmalarda sınıflandırma performansının arttığı görülmüştür. Daha sonra bulunan öznelik oluşturma yöntemlerinin başarımını test etmek için ele alınan veri kümesi dışındaki, uyku apnesi ve fiziksel aktivite veri kümelerinde de sınıflandırma performansı test edilmiştir. Farklı veri kümeleri ile kullanılan yeni öznelikler ile yapılan çalışmalarda sınıflandırma performansının arttığı görülmüştür. Geleneksel sınıflandırma yöntemleri ile beraber bu tez çalışması kapsamında geliştirdiğimiz CNN+LSTM mimarisi karşılaştırılmış; sınıflandırma analizlerinde CNN+LSTM mimarisi ile performans artışı sağlandığı tespit edilmiştir.

**ANAHTAR SÖZCÜKLER:** standart sapma ve üstel hareketli ortalama ile öznelik, derin öğrenme, uyku apnesi, açık ve kapalı hava ile fiziksel aktivite sınıflandırma.

**Danışman:** Dr. Öğr. Üyesi Mustafa SERT, Başkent Üniversitesi, Bilgisayar Mühendisliği Bölümü.

## **ABSTRACT**

### **MULTIMODAL AND FUSION BASED ANALYSIS OF WEARABLE SENSOR DATA FOR HUMAN HEALTH MONITORING**

Gökhan MEMİŞ

Başkent University Institute of Science and Engineering

**Department of Computer Engineering**

In this study, in the field of health informatics, intelligent methods based on multi-mode and fusion-based analyzes have been developed in order to obtain meaningful information from the data collected from different sensors in order to monitor the health status and daily activities of the elderly and people in need of care. Pulse, EKG, oxygen saturation is used as wearable sensors. In addition to arrhythmia, sleep apnea diseases and early diagnosis of the data collected from these sensors, the physical activity of the person who was monitored by accelerometers, magnetometer, gyroscope and inclinometer sensors was investigated. In the scope of the thesis, physical activity data were collected from 12 different people according to the presence of indoor and outdoor. With the data collected, it was found where the person's physical activity was done. Standard deviations and exponential moving average based new features are created at this stage. It is observed that the performance of classification has increased in the studies conducted with the new features. Subsequent to our own dataset, the performance of the classification was also tested in sleep apnea and physical activity datasets to test the performance of the attribute building methods found later. It has been observed that the performance of classification has increased in studies with new attributes used with different datasets. The CNN + LSTM architecture we developed within the scope of this thesis was compared with the traditional classification methods. In the classification analysis, it has been found that performance increase is achieved with CNN + LSTM architecture.

**KEYWORDS:** features with standard deviation and exponential moving average, deep learning, sleep apnea, physical activity with indoor and outdoor information.

**Adviser:** Asst. Prof. Dr. Mustafa SERT, Başkent University, Department of Computer Engineering.



# İÇİNDEKİLER LİSTESİ

Sayfa

ÖZ .....	i
ABSTRACT .....	ii
İÇİNDEKİLER LİSTESİ .....	iii
ŞEKİLLER LİSTESİ .....	vi
ÇİZELGELER LİSTESİ .....	viii
SİMGELER VE KISALTMALAR LİSTESİ .....	x
1. GİRİŞ .....	1
1.1. Problem Tanımı .....	2
1.2. Tezin Amacı, Konu ve Kapsamı .....	3
1.3. Araştırma Soruları .....	4
1.4. Tezin Önemi .....	4
2. ÖNCEKİ ÇALIŞMALAR .....	7
2.1. Fiziksel Aktivite Tanıma .....	7
2.2. Uyku Apnesi ve Kalp Krizi Tanıma .....	10
3. KULLANILAN TEMEL YÖNTEMLER .....	15
3.1. Sınıflandırma Algoritmaları .....	15
3.1.1. K En Yakın Komşu (kNN) Algoritması .....	15
3.1.2. Naive (Saf) Bayes Sınıflandırma Algoritması .....	16
3.1.3. Destek Vektör Makineleri (SVM) .....	17
3.1.4. Rastgele Orman (Random Forest) .....	19
3.1.5. Yapay Sinir Ağları (YSA) .....	21
3.1.6. Çok Katmanlı Perceptron (Multilayer Perceptiron (MLP)) .....	24
3.1.7. Evrimsel Sinir Ağı (ConvNet / CNN) .....	24
3.1.8. Uzun / Kısa Süreli Bellek (Long / Short Term Memory (LSTM)) .....	26
3.2. Öznitelik Seçme Yöntemleri .....	27
3.2.1. Relieff Öznitelik Seçme Algoritması .....	27

3.2.2.	Ki-Kare Öznitelik Seçme Algoritması .....	27
3.2.3.	Bilgi Kazancı Öznitelik Seçme Algoritması .....	28
3.2.4.	Kazanım Oranı Öznitelik Seçme Algoritması .....	29
4.	ORTAM BAZLI FİZİKSEL AKTİVİTE TANIMA İÇİN VERİ TOPLAMA .....	30
4.1.	Donanımların Kurulumu .....	30
4.2.	Akıllı Telefon Sensörleri .....	35
4.3.	Denekler .....	36
4.4.	Veri Toplama Protokolü .....	36
5.	SENSÖR ANALİZLERİ .....	39
5.1.	Öznitelik Seçimi Yöntemi .....	40
5.2.	Öznitelik Füzyonu .....	41
5.3.	Sınıflandırıcı Tasarımı .....	42
5.4.	Kullanılan Veri Kümeleri .....	43
5.4.1.	HPAwIO Veri Kümesi .....	43
5.4.2.	MHealth Veri Kümesi .....	43
5.4.3.	Physionet Sleep Apne Veri Kümesi .....	43
5.4.4.	Statlog Veri Kümesi .....	44
5.4.5.	MIT-BIH Veri Kümesi .....	44
5.5.	DeneySEL Çalışmalar .....	45
5.5.1.	HPAwIO Veri Kümesi Çalışmaları .....	45
5.5.2.	MHealth Veri Kümesi Çalışmaları .....	48
5.5.3.	Physionet Veri Kümesi Çalışmaları .....	51
5.5.4.	Statlog Veri Kümesi Çalışmaları .....	61
5.5.5.	MIT-BIH Veri Kümesi Çalışmaları .....	61
6.	GÜRBÜZ ÖZNİTELİK ÇIKARIMI YÖNTEMLERİ .....	63
6.1.	Sigma Tabanlı Öznitelik Oluşturma .....	63
6.2.	Üstel Hareketli Ortalama Tabanlı Öznitelik Oluşturma .....	64
6.2.1.	Tekil üstel hareketli ortalama (EMA) .....	66

6.2.2.	Çiftli üstel hareketli ortalama (DEMA) .....	67
6.2.3.	Üçlü üstel hareketli ortalama (TEMA) .....	68
7.	SINIFLANDIRICI TASARIMI .....	69
7.1.	Derin Yapay Sinir Ağları (DNN).....	69
7.2.	Evrışimsel Sinir Ağı (ConvNet / CNN) .....	72
7.3.	Uzun / Kısa Süreli Bellek (Long / Short Term Memory (LSTM)).....	73
7.4.	Hibrid CNN+LSTM Algoritması.....	73
8.	DENEYSEL BULGULAR.....	76
8.1.	HPAwIO Veri Kümesi Bulguları .....	76
8.1.1.	Senaryo bazlı sınıflandırma bulguları .....	79
8.1.2.	Sigma tabanlı öznitelik ile sınıflandırma bulguları .....	81
8.1.3.	EMA öznitelikler kullanıldığında farklı veri kümelerindeki bulguları ...	85
8.2.	Physionet Veri Kümesi Bulguları .....	86
8.2.1.	Sigma tabanlı öznitelik ile sınıflandırma bulguları .....	86
9.	TARTIŞMA VE SONUÇ .....	88
	KAYNAKLAR.....	91

## ŞEKİLLER LİSTESİ

Şekil 3.1 k En Yakın Methodu Örneği.....	16
Şekil 3.2 Karar ağacı örneği .....	20
Şekil 3.3 İkili adım aktivasyon fonksiyonu gösterimi .....	22
Şekil 3.4 Sigmoid aktivasyon fonksiyonu gösterimi .....	22
Şekil 3.5 Relu aktivasyon fonksiyonu gösterimi.....	23
Şekil 3.6 Tanh ile Sigmoid aktivasyon fonksiyonları arasındaki fark.....	23
Şekil 3.7 Evrişimsel Sınır Ağı Örneği .....	25
Şekil 4.1 Kullanılan Oksimetre Cihazı .....	31
Şekil 4.2 Cep telefonu eksenleri .....	31
Şekil 4.3:Veri toplama mimarisi .....	33
Şekil 5.1 Öznitelik füzyonu .....	41
Şekil 5.2 Farklı yaş grupları için kullanılan yöntemlerin, kişinin açık ya da kapalı alanda olup olmadığının tanınma doğruluğu. ....	46
Şekil 5.3 Senaryo bazlı fiziksel aktivite sınıflandırması için makine öğrenme yöntemlerinin ortalama doğrulukları .....	50
Şekil 5.4 WSwSS test senaryosu için tek ve çok kipli özniteliklerin uyku apnesi sınıflandırması için performans karşılaştırması .....	53
Şekil 5.5 BSwSS test senaryosu için tek ve çok kipli özniteliklerin uyku apnesi sınıflandırması için performans karşılaştırması .....	54
Şekil 5.6 BSwDS test senaryosu için tek ve çok kipli özniteliklerin uyku apnesi sınıflandırması için performans karşılaştırması .....	54
Şekil 5.7 Relieff ile seçilen öznitelikler kullanıldığında uyku apnesi için DVM doğruluk performansı .....	59
Şekil 5.8 Relieff ile seçilen öznitelikler kullanıldığında uyku apnesi için NB doğruluk performansı .....	59
Şekil 5.9 Relieff ile seçilen öznitelikler kullanıldığında uyku apnesi için kNN doğruluk performansı (k=3) .....	60
Şekil 5.10 Kalp hastalıkları için farklı sınıflandırma algoritmalarına göre doğruluk değerlerinin karşılaştırması .....	61
Şekil 5.11 Aritmi için sınıflandırma algoritmalarına göre doğruluk değerleri .....	62
Şekil 5.12 Aritmi için sınıflandırma algoritmalarına kişi bazlı doğruluk değerleri... ..	62
Şekil 6.1 Ortalama $\pm$ Sigma Gösterimi Örneği.....	64

Şekil 6.2 Eđimölçer verisi kesiti için SMA ile EMA karşılaştırması.....	65
Şekil 6.3 Magnetometre verisi kesiti için DEMA ile EMA karşılaştırması .....	67
Şekil 6.4 Magnetometre verisi kesiti için TEMA ile DEMA ve EMA karşılaştırması .....	68
Şekil 7.1 Derin Yapay Sinir Ađı Mimarisi .....	70
Şekil 7.2 Kullanmış olduğumuz CNN mimarisi .....	72
Şekil 7.3 Kullanmış olduğumuz LSTM mimarisi.....	73
Şekil 7.4 Kullanmış olduğumuz CNN + LSTM mimarisi.....	75
Şekil 8.1 Açık ile kapalı alanlar arasında SpO <sub>2</sub> deđişimi.....	78
Şekil 8.2 Sigma tabanlı öznitelik kullanarak veri kümelerinin karşılaştırılması .....	84
Şekil 8.3 Uyku apnesi sınıflandırması için sensor verisi, öznitelik seçme yöntemi ile dođruluk karşılaştırması .....	87

## ÇİZELGELER LİSTESİ

Çizelge 2.1 Fiziksel Aktivite Tanıma Çalışmaları.....	10
Çizelge 2.2 Uyku Apnesi Tanıma Çalışmaları .....	14
Çizelge 4.1 Sensörlerin vücutta yerleşme yerleri ve adetleri (örn. Elde tutulan telefonda 3 adet ACC sensörü vardır) .....	33
Çizelge 4.2 Tanımlanan Fiziksel Aktivite Sınıfları .....	34
Çizelge 4.3 Toplanan sensör verilerinin bazı istatistiksel değerleri.....	34
Çizelge 4.4 Deneklerin özellikleri.....	36
Çizelge 4.5 Deneklerin yürüme ve koşma hızları .....	38
Çizelge 4.6 Karşılaştırma için kullanılan veri kümelerinin karakteristikleri .....	38
Çizelge 5.1 Öznitelik seçimine göre füzyon edilmiş ve tek kipli olarak oluşturulan senaryolar.....	40
Çizelge 5.2 Farklı yaş grupları için istatistik değerleri.....	46
Çizelge 5.3 Fiziksel aktivite sınıflandırma metodlarının karşılaştırması.....	48
Çizelge 5.4 Verilerin tek kipli şeklinde fiziksel aktivite tahminindeki başarımına etkisi (0: en düşük, 1: en yüksek) .....	49
Çizelge 5.5 Senaryo (Sen) bazlı fiziksel aktivite sınıflandırması için makine öğrenme yöntemlerinin doğruluklarının karşılaştırma tablosu .....	49
Çizelge 5.6 Kullanılan method ile [88] ve [89] yöntemleri arasındaki başarımların karşılaştırması .....	51
Çizelge 5.7 Çapraz doğrulama fold (katlama [kat]) karşılaştırması .....	53
Çizelge 5.8 NB, SVM ve kNN modellerinin uyku apnesi sınıflandırması için performans karşılaştırması .....	56
Çizelge 5.9 Metodumuzun [96] 'da olan çalışmalarla doğruluk karşılaştırması ....	56
Çizelge 5.10 Metodumuzun [97] 'da olan çalışmalarla doğruluk karşılaştırması ..	56
Çizelge 5.11 Kullanılan öznitelik seçme methodu ile bulunan özniteliklerin sayılarına göre uyku apnesi sınıflandırma için makine öğrenme başarımları (*: Öznitelik seçimi yapılmadan).....	57
Çizelge 5.12 Kullanılan yöntemimiz ile [96] ve [97] çalışmalarında kullanılan en yüksek doğruluğu veren methodların karşılaştırması. ....	60
Çizelge 7.1 Derin yapay sinir ağı için Deeplearnin4j kütüphanesi parametreleri ..	71
Çizelge 7.2 Evrişimsel sinir ağı için Deeplearnin4j kütüphanesi parametreleri .....	72

Çizelge 7.3 Uzun/Kısa Süreli Bellek Sinir Ağı için Deeplearnin4j kütüphanesi parametreleri .....	73
Çizelge 7.4 Hibrid CNN+LSTM Sinir Ağı için Deeplearnin4j kütüphanesi parametreleri .....	75
Çizelge 8.1 Tek kipli ile çok kipli veriler için örneklem oranlarına göre sensör FA sınıflandırma performansları (doğrulukları).....	77
Çizelge 8.2 Oksimetre sensörü kullanma esnasında açık ya da kapalı alan sınıflandırma performansları (A: 18-34 yaş aralığı, B: 35-54 yaş aralığı, C: 56-65 yaş aralığı).....	78
Çizelge 8.3 Açık ya da kapalı alan içeren FA için sınıflandırma methodlarını karşılaştırması .....	79
Çizelge 8.4 Açık ya da kapalı alan içeren FA için önerilen öğrenme algoritmaları ile sonuçlar.....	81
Çizelge 8.5 Farklı yaş ve kişiye göre önerilen yöntemin açık ya da kapalı alan içeren FA için doğruluklarının karşılaştırılması .....	82
Çizelge 8.6 Önerilen Sigma tabanlı öznitelikler ile DNN kullanıldığında ortalama precision ve recall değerleri (%) .....	83
Çizelge 8.7 Açık ya da kapalı alan içeren FA sınıflandırma için sigma tabanlı öznitelikleri (özn.) kullanarak DNN ve CNN+LSTM karşılaştırması .....	84
Çizelge 8.8 Açık ya da kapalı alan içeren FA sınıflandırması için üstel hareketli ortalama tabanlı öznitelikleri kullanarak veri kümelerinin karşılaştırılması.....	85

## **SİMGELER VE KISALTMALAR LİSTESİ**

ACC	Accelerometer
AHI	Apne Hipopne Index
CLI	Clinometer
CV	Cross Validation
DEMA	Double Exponential Moving Average
DNN	Deep Neural Network
ECG	Elektrokardiyogram
EEG	Elektroensefalogram
EMA	Exponential Moving Average
EMG	Elektromiyogram
EOG	Elektrookülogram
GYR	Gyroscope
kNN	K Nearest Neighborhood
MAG	Magnetometer
FA	Physical Activity
PSG	Polisomnografi
RBF	Radial Basis Function (SVM kernel)
STS	Short Sleeved
SVM	Support Vector Machine
TEMA	Triple Exponential Moving Average



## 1. GİRİŞ

Mobil cihazların ve taşınabilir bilgisayarların gelişmesi hayatımızı çok hızlı bir şekilde değiştirmektedir. Özellikle, son yıllarda giyilebilir teknolojilerin kullanımı her alanda (sağlık bilişimi, akıllı ev uygulamaları, savunma vb.) artmaktadır. Bu teknolojiler, tıpkı bir bilezik ya da aksesuar gibi taşınarak, içinde bulunan belli sensörler sayesinde farklı görevleri yerine getirebilmektedir. Bu sensörlerin topladığı veriler bluetooth ya da benzeri kablosuz ağlar aracılığıyla veritabanlarına ya da kişisel bilgisayarlara gönderilebilmektedir [1]. Sensör teknolojilerindeki bu gelişim, bilgisayar ve işlemci boyutlarının küçülmesi ile birlikte, giyilebilir teknolojilerin sağlık alanında nasıl kullanılabileceği konusu yoğun olarak araştırılmaya başlanmıştır.

Giyilebilir cihazlar, sınırları içine başka konuları veya anlamları içermesinden dolayı çok sayıda disiplin barındırmaktadır ve farklı disiplinlerden birçok araştırmacıyı bu alanda çalışmalar yapmaya yönlendirmektedir. Hastalıklar, kronik rahatsızlıklar, gündelik aktivite takibi gibi çok çeşitli sebepler için fizyolojik parametrelerin ölçülmesi amacıyla tasarlanan bu cihazlar, özellikle sağlık alanında son kullanıcıların hizmetine sunulmaktadır. Bununla birlikte sağlık, spor, eğlence, eğitim, iş ve daha birçok alanda hayatlarımızda giderek daha fazla yer bulmaya başlayan giyilebilir sensörler ile pek çok ölçüm yapılabilmektedir. Farklı donanım ve kullanım yöntemlerine göre sınıflandırılan giyilebilir sensörlerin ortak özellikleri arasında, yüksek kapasitede kablosuz iletişim yapabilme ve dâhili ölçüm yapan sensörleri barındırma yer almaktadır.

Bu çalışmada, sağlık bilişimi alanında, yaşlı ve bakıma muhtaç kişilerin, sağlık durumlarının ve günlük aktivitelerinin takibi amacıyla, farklı sensörlerden toplanan verilerden anlamlı bilgilerin elde edilebilmesi için, çok kipli ve füzyon-tabanlı analizlere dayalı akıllı yöntemlerin geliştirilmesi hedeflenmektedir. Giyilebilir sensörler olarak, Elektrokardiyografi (EKG), nabız, oksijen doygunluğu kullanılmıştır. Bu sensörlerden toplanan veriler ile aritmi, uyku apnesi ve kalp krizi gibi hastalıkların takibi ve erken tanılmanın yanı sıra, ivme ve eğim sensörleri ile izlenen kişinin fiziksel aktivitesi incelenmiştir. Tez çalışması kapsamında standart sapma tabanlı öznitelik oluşturma, üstel hareketli ortalama tabanlı öznitelik oluşturma, veri füzyonu, ilinti, bağlanım analizi, makine öğrenme algoritmaları ve derin öğrenme algoritmaları kullanılmıştır.

## 1.1. Problem Tanımı

Dünyadaki yaşlı nüfusunun varlığı, yaşlıların çoğunlukla yardıma ihtiyaç duyması insanları yaşlı bakımı problemi ile yüz yüze getirmektedir. Evde bakım, yaşlılar bakımından bazı cazip yönleri olmasına rağmen, sorunsuz değildir. Sorunlar daha çok, yaşlı ebeveynleri ve evde onların bakım sorumluluğunu üstlenen yetişkinleri ilgilendirmektedir. Devamlı olarak yaşlının yanında bulunmak ve sağlık durumunu kontrol etmekte bakıcı kişiler için de sorun oluşturmaktadır. Yaşlı hastaların takibi bir hekim ve yaşlı hastanın ailesi için zorunlu ve önemli aktivitelerin başında gelir. Yaşlı nüfusun kalabalıklaşması ile beraber artan sağlık ve bakım maliyetleri, sağlık sisteminin temelinde bazı köklü değişikliklere gidilmesinin ihtiyacını doğurmaktadır. Yaşlanan dünya nüfusu ile artan sağlık maliyetleri ev tabanlı tarama testlerini ve yaşlıların evinde izlenmesi teknolojileri gibi etkili yöntemlere ihtiyaç artmaktadır.

Yukarıda bahsedilen yaşlı sayısındaki artış doğrultusunda bazı nedensel sorular ile karşılaşılmaktadır. Bunlar,

- Karmaşık tıbbi koşullar altında yaşlı kişiler ile nasıl ilgilenebiliriz?
- Bakıcılar ya da kendi çocukları tarafından erişimi zor olan yerlerde nasıl yaşlı bakımını yaparız?
- Sağlık problemi olan yaşlıların bakımında, teknoloji kullanmak hayata tutunmalarını nasıl maksimize eder?
- Yaşlı bakımında meydana gelen sorunların temel sebebi nedir?
- Bu sorunların içerisinde teknoloji kullanılarak aşılabilecek sorunlar var mıdır?
- Teknoloji kullanılarak aşılabilecek sorunlara ilişkin çözüm yöntemleri neler olabilir?

olarak sıralanabilir. Bu sorular yavaş yavaş çözülmeye başlamasına rağmen, gelecek için de daha çok yol alınması gerekmektedir. Alınacak yollar ile beraber maliyetler de artacaktır. Bu çalışmada da çözümün bir parçası olarak bilişim teknolojileri kullanılacaktır. Şu anda da bilişim teknolojileri kullanılarak yaşlıların takibini yapan uygulamaların geliştirilmesi artarak devam etmektedir [1]. Bununla birlikte Pricewaterhouse Coopers Health Research Institute'nün (PwC) 2014 yılında Amerika'da yapmış olduğu araştırmaya göre katılımcıların %56'sı on sene içerisinde giyilebilir sensörlerin kullanımına bağlı olarak ortalama yaşam süresinin artacağını düşünmektedir [2].

Giyilebilir sensörlerin hem tanısal hem de izleme uygulamaları vardır. Bunların mevcut yapabildikleri hem fizyolojik ve biyokimyasal algılama hem de hareket algılama içerir [3]. Bu teknolojilerin çözmeye yardımcı olabileceği sorunların sayısı daha da fazla olabilir. Örneğin, psikolojik izleme, teşhis ve psikolojik olan nöbetler, hipertansiyon, ritim bozukluğu ve astım gibi kardiyovasküler ve akciğer hastalıkları olan bireylerin tedavisinde ve takibinde yardımcı olabilir.

Akıllı telefonlar ve kablosuz bağlantıların yaygın kullanımı kişilerin günlük yaşantısını devam ettirmeleri esnasında gerçek zamanlı izlemeyi mümkün kılmaktadır. Uzaktan takip teknolojileri gerçek zamanlı sağlık bilgilerini sağlama potansiyeline sahip ve aynı zamanda kendi evlerinin konforunda hastalar için önemli bilgilerin takibi için bir araçtır [4].

Kentsel alanlara kıyasla kırsal kesimde olanlar için herhangi bir aksi durumda ulaşım ve tedavi zor olduğu için kötü koşulların ortaya çıkma ihtimali daha yüksektir. Doktora ya da herhangi bir yakınına ulaşmak şehirde yaşayanlara göre çok daha zordur. Böyle durumda olanlarda kalp krizi ya da şeker koması sıkıntıları çok daha tehlikeli sonuçlara sebep olmaktadır [5].

Yukarıda bahsedilenler doğrultusunda, durmadan değişen sağlık alanında, farklı sistemleri ve cihazları birleştirme, bilişim teknolojileri ile insanlara gerçek faydalar sağlamaktadır. Bilişim teknolojileri ile birlikte çalışabilirliğe ulaşmak, rekabetçi olmak ve hasta bakımını sürekli iyileştirmek, sağlık kurumlarının artan hedefidir. Bu teknolojiler, hasta yatağının yanında, ameliyathanede, eczanede ve daha fazlasında, anlamlı veriye daha iyi erişim sağlamak adına veri akışlarını entegre eden ve kullanan çözümler yaratan Nesnelerin İnterneti'nin (Internet of Things – IoT) bir parçasıdır.

Yaşlılar için giyilebilir sensörler ile ilgili çalışmalar son yıllarda hız kazanmıştır. Yapılan çalışmalar ile hastanın nabız, tansiyon, ateş durumu, kalp durumu gibi bilgileri takip edilmekte, bunlarla beraber bazı hastalıkların erken teşhisine yardımcı olmaktadır [84].

## **1.2. Tezin Amacı, Konu ve Kapsamı**

Sağlık bilişimi alanında, yaşlı ve bakıma muhtaç kişilerin, sağlık durumlarının ve günlük aktivitelerinin takibi amacıyla, farklı sensörlerden toplanan verilerden anlamlı bilgilerin elde edilebilmesi için, çok kipli ve füzyon tabanlı analizlere dayalı akıllı

yöntemler geliřtirmek hedeflenmiřtir. Giyilebilir sensörler olarak nabız, EKG, oksijen doygunluęu yanı sıra, cep telefonu içerisinde bulunan sensörlerden ivmeölçer, manyetik alan ölçer, jiroskop ve eęimölçer sensörleri ile izlenen kiřinin aritmi, uyku apnesi ve fiziksel aktivitesinin açık ya da kapalı alanda yapılıp yapılmadıęının bulunması amaçlanmıřtır. Tez çalıřması kapsamında, açık ve kapalı alan bilgisi ile birlikte fiziksel aktivite veri kümesi oluřturma, gürbüz öznitelik analizleri ve derin öğrenme modelleri ile mevcut başarımların artırılması hedeflenmiřtir. Ayrıca öznitelik oluřturma yöntemlerinin başarımlarını test etmek için bu tez çalıřması kapsamında oluřturulan veri kümesi dıřındaki, uyku apnesi ve fiziksel aktivite veri kümelerinde de sınıflandırma performansı test edilmesi hedeflenmiřtir.

### **1.3. Arařtırma Soruları**

Bu arařtırmada ařaęıdaki arařtırma soruları ve varsayımlardan hareket edilmiřtir;

- Giyilebilir sensörler ile elde edilen verilerden yola çıkılarak hastalık takip, teřhis, sınıflandırma gibi analizlerin yapılması mümkün müdür?
- Tek giyilebilir sensörden elde edilecek veriler yerine birden fazla giyilebilir sensörden elde edilecek veriler (çok kipli) hastalık takibi ve teřhise yardımcı olmak için daha verimli sonuçlar verebilir mi?
- Birden fazla sensörden gelen verileri doğrudan alıp analiz etmek yerine veri füzyonu yapılıp analiz etmek sınıflandırmanın başarımlarını etkiler mi?
- Bir kiřinin fiziksel aktiviteyi açık ya da kapalı alanda yaptıęı bilgisini bulabileceęimiz bir sensör var mıdır?
- Sensör verileri yerine zaman serisi bazlı öznitelikleri kullanmak sınıflandırma başarımlarını etkiler mi?

### **1.4. Tezin Önemi**

Mevcut literatürde görüldüęü üzere genelde sensörlerden gelen veriler tek kip analiz ile tek durumu takip ya da tahmin etmek için kullanılmıřtır. Örneęin fiziksel aktivite sınıflandırmasının yanında fiziksel aktiviteyi açık ya da kapalı alanda mı yapıldıęı bilgisini bulan bir çalıřma yoktur. Tek sensörden gelen verilere göre sınıflandırma yapan çalıřmalar bulunmaktadır. Ayrıca, birden fazla sensör kullanıldıęında daha ayrıntılı bilgileri bulmak ve sınıflandırma başarımlarını arttırmak mümkündür.

Bu çalışmada, veri toplamak ve fiziksel aktivite tanıma işlemi gerçekleştirmek için akıllı telefonlar ile oksimetre sensörü kullanılmıştır ve deneklerin açık ya da kapalı alanda fiziksel aktivite yapıp yapmadıklarını keşfetmek amaçlanmıştır. Çalışma kapsamında yürüme, koşma, ayakta durma ve oturma gibi temel fiziksel aktiviteler seçilmiştir ve bu aktiviteler hem açık hem de kapalı alanlarda gerçekleştirilmiştir. Ek olarak, bir aktivitenin zamansal özelliklerini daha iyi yakalamak, böylece tanıma doğruluğunu iyileştirmek için sigma tabanlı ve üstel ortalama tabanlı öznitelikler sunulmuştur.

Tez çalışmasının özgün katkılarının aşağıdaki kapsamda olacağı değerlendirilmektedir:

- İvmeölçer, manyetometre, jiroskop, klinometre ve kullanıcının cep telefonunun ortam sıcaklığı verileri ile SpO<sub>2</sub> sensör verilerinden oluşan yeni, çok kipli veri kümesi oluşturulmuştur. 12 farklı yaş grubundaki gerçek kişiler ile toplamda, 16 saat kayıt vardır ve açık ya da kapalı alan durumlarına göre dört farklı aktivite için etiketlenmiştir.
- Birden fazla sensör kullanmak veri kümesindeki aykırı değerleri artırabilir; ancak, veri içerisinde uç değerlerin tanımlanmasının, her veri kümesinin özellikleri nedeniyle kesin bir yolu yoktur. Bunun yerine, ham gözlemler yorumlanmalı ve bir değer in uç değer olup olmadığına karar verilmelidir. Bu amaçla, sınıflandırma öncesinde öznitelik oluşturmak ve daha önce bahsedilmemiş olan tanıma için üstel hareketli ortalama ile standart sapma tabanlı öznitelikler sunulmuştur.
- Veri kümelerindeki sınıflandırma görevleri için yeni derin sinir ağı (DNN) mimarileri tasarlanmıştır ve verimliliğini göstermek için kapsamlı analizler yapılmıştır.
- Oksimetre sinyalinin açık ve kapalı alanlarda nasıl etkilendiğine dair analizler yapılmıştır. Ayrıca, çoklu sensör kullanımının yanı sıra farklı öğrenme yöntemleriyle birlikte, açık ve kapalı alan bilgisi içeren fiziksel aktivitenin tespit doğruluğu üzerindeki etkisini araştırılmıştır.
- Çok kipli karar modelleri (füzyon) oluşturulmuştur,
- Tek kipli ve çok kipli modellerin ele alınan veri kümesi ve problemdeki karar verme başarımlarının karşılaştırılması yapılmıştır,

Çalışma sonucunda elde edilecek bulguların, sağlık bilişimi alanındaki farklı araştırmalara ışık tutacağı değerlendirilmektedir.

Bu çalışmanın geri kalanı şu şekilde düzenlenmiştir: 2. Bölüm fiziksel aktivite, uyku apnesi ve kalp hastalıkları ile alakalı geçmiş çalışmaları özetlemektedir. 3. Bölüm tez süresinde kullanılan temel yöntemleri açıklamaktadır. 4. Bölüm tez çalışması kapsamında toplanılan veri kümesi ile alakalı bilgiler vermektedir. 5. Bölüm kullanılan sensörlerin analizlerini sunmaktadır. 6. Bölüm gürbüz öznelik çıkarma yöntemlerini açıklamaktadır. 7. Bölüm kullanılan özgün derin öğrenme algoritmalarını açıklamaktadır. 8. Bölüm bulunan deneysel bulguları içermektedir. Son bölüm ise tartışma ve sonuçlardır.

## 2. ÖNCEKİ ÇALIŞMALAR

### 2.1. Fiziksel Aktivite Tanıma

İnsan fiziksel aktivitesini (Fiziksel Aktivite [FA]) otomatik olarak tanımak, akıllı evler, sağlık izleme ve acil servisler gibi farklı uygulamalarda esastır. Uzun süreli sağlık izleme sistemlerini kullanmak için, insan faaliyetlerini otomatik olarak tanımlayan akıllı cihazlar gerekir. İnsan fiziksel aktivitelerinin tanınması sağlık hizmetleri ve akıllı ortamlar için çok önemli bir araştırma alanıdır [6]. Kalabalık kentsel ortamlar, yoğun trafik, hava kirliliği, spor için yeşil alan eksikliği, sokaklardaki suç korkusu, kışın soğuk hava koşulları ve daha az gün ışığından yararlanma, insanların dışarıda egzersiz yapmasını engelleyebilir. Bu nedenlerden dolayı FA için spor salonları veya evler tercih edilebilir. Üstelik yeşil bir alanda yapılan FA'nin hemen sonrasında daha fazla mutluluk oluşur [7]. Bu nedenle FA'nin nerede yapıldığını bilmek gerekebilir. Daha önceki bazı çalışmalar ile açık ya da kapalı alan algılamaya odaklanılarak temel bilgiler toplanmıştır [7], [8]. Bu çalışmalar genellikle bu bilgiyi elde etmek için bir cep telefonu veya GPS cihazı kullanır [9]. Ancak, cep telefonu sinyali servis sağlayıcıya ulaşmadığında veya GPS sinyali olmadığında açık ve kapalı alan bilgilerini bulunamayabilir.

Öte yandan, sağlık hizmeti izleme sistemlerinde, oksijen satürasyonu ( $SpO_2$ ), vücut ısısı, kalp atış hızı ve kan basıncı gibi fizyolojik sensörler birçok uygulamada yaygın olarak kullanılmaktadır. Bunlar arasında  $SpO_2$ , hipoksemi, grip, astım atakları, kronik obstrüktif akciğer hastalığı, uyku apnesi ve kardiyovasküler hastalıklar gibi birçok amaç için kullanılmıştır [10]. Bilindiği kadarıyla, oksimetre sensörü literatürde FA'nin açık ya da kapalı alanda yapılıp yapılmadığını bulmak için kullanılmamıştır. Literatürde, vücut sensörleri ve akıllı telefonlar popüler hale geldiğinden beri insan faaliyetlerini sınıflandırmak için cep telefonlarının kullanıldığı anlaşılmaktadır [11]. Bu cihazlar FA'leri ve günlük rutinleri tanımlamak için daha fazla bilgi sağlar ve kullanıcılara özgü parametreleri ölçmek için mükemmel bir başlangıç noktasıdır [12]. Oksijen seviyesi FA veya çevresel koşullara göre de değişebilir. Bu nedenlerden dolayı,  $SpO_2$ 'nin sürekli izlenmesi, akciğer veya kalp ile ilgili hastalıkları olan kişiler için olduğu gibi FA'larını izlemek isteyenler için de önemlidir. Tüm bunlara dayanarak,  $SpO_2$  sinyalinin açık ya da kapalı alanda nasıl etkilendiğini ve bu fizyolojik sinyalin kişilerin açık ya da kapalı alanda FA ile meşgul olup olmadığını bulmak için kullanılıp kullanılmayacağı araştırılmıştır.

Giyilebilir sensörler fiziksel bilgileri algılayabilir, toplayabilir ve yükleyebilir. Dahası, bir cep telefonu ile elde edemediğimiz bilgileri verebilir [13]. FA tespit uygulamaları sağlıklı bir yaşam tarzını destekleyebilir. Öte yandan, yaşlıların ve küçük çocukların düşmesi veya kaybolması gibi istenmeyen durumları önleyebilirler [14]. Bu tür bir uygulama, insanların ne kadar yürüdüğü, kaç adım attıkları, ne kadar uydukları ve kaç kalori yaktıkları gibi yararlı veriler sağlayabilir.

Geleneksel örüntü tanıma yaklaşımları FA'nın tanınmasında çok başarılı olmuştur [15]. Ancak, bu yaklaşımların bazı dezavantajları vardır. Birincisi, öznitelikler, alan hakkında bilgi sahibi olan kişilerin yardımıyla çıkarılır. Bu insan bilgisi basit, iyi bilinen FA'yı tanımaya yardımcı olur; ancak FA karmaşıksa veya çevre ile birlikte FA'nın tanınması gerekiyorsa, işler daha zor ve zaman alıcı hale gelir. İkincisi, insanlar yalnızca alan bilgisiyle sınırlı olan öznitelikleri öğrenebilirler [16]. Bu yetersiz öznitelikler genel olarak ortalama, standart sapma ve sıklık gibi bazı istatistiksel bilgileri ifade eder. Yürüme veya koşma gibi basit etkinlikleri tanımak için kullanılabilirler, ancak karmaşık etkinlikleri veya etkinlik ortamını ortaya çıkarmakta zorlanabilirler [17]. Örneğin, açık alan yürüyüş aktivitesi daha karmaşık ve sadece basit özellikler kullanılarak tanınması neredeyse imkânsızdır.

Derin öğrenme, bu sınırlamalar karşısında başarılı sonuçlar verme eğilimindedir. Bu nedenlerden dolayı, derin öğrenme FA tanıma için ideal bir yaklaşımdır ve mevcut literatürde geniş ölçüde araştırılmıştır [18 - 20]. Jun vd. [18] geleneksel UCI HAR veri kümesi ve bunların zayıf etiketli veri kümesi olmak üzere iki farklı veri kümesi kullanmışlardır. Tekrarlayan dikkat öğrenme (attention learning) modelini değerlendirmişlerdir. Ordóñez ve Roggen [19], entegre evrişimli ve tekrarlayan katmanlar kullanarak DeepConvLSTM'yi önermiştir. Ha ve Choi [20], mekânsal ve zamansal faaliyet ilişkilerinin derin temsilinden dolayı geleneksel yaklaşımların doğruluğunu geliştirmek için Konvolüsyonel Sinir Ağı geliştirmiştir. Bu tür sinir ağları görüntüyü kullanarak tahminlerde bulunur ve ardından görüntüde belirli bileşenlerin olup olmadığını tespit eder. Bunlar hesaplama açısından pahalıdır. Bu nedenle, görüntüye dönüştürülmeden sadece sensör verilerinin kullanılması avantajlı olabilir.

Valentin vd. [21], modern akıllı telefonlarda kolayca bulunabilen düşük güç sensörleri kullanarak bir kullanıcının açık alanda mı ya da kapalı alanda mı olduğunu belirleme problemini ele almışlardır. Kullanıcının içeride mi yoksa dışarıda mı



olduğunu bulmak için ışık, manyetik alan ve bir cep telefonunun hücresel sinyal detektörü kullanmışlardır. Bununla birlikte, hafif ve manyetik bir alan detektörü kullanmak zordur ve ayrıca, cep telefonu sinyali servis sağlayıcıya ulaşmadığında veya GPS sinyali olmadığında açık ve kapalı alan konum bilgilerini bulunamayabilir.

İvmeölçer (ACC – Accelerometer) sensörleri, FA sınıflandırması dahil olmak üzere çeşitli uygulamalarda kullanılmıştır. Wang vd. [22], EMD (Ampirik Mod Ayırma [Empirical Mode Decomposition]) yöntemini kullanarak giyilebilir üç eksenli ACC'den hareket verilerinin özelliklerini çıkarmışlardır. Sensörler bel ve ayak bileği bölgelerine takılmıştır. Oyun teorisini kullanarak öznitelikleri seçmişlerdir. Ayrıca SVM (Destek Vektör Makinaları [Support Vector Machine]) ve kNN (k En Yakın Komşu [k nearest neighbor]) ile sınıflandırırılar. 9 farklı FA için genel olarak % 80 tanıma doğruluğu rapor etmişlerdir. Mingqi vd. [23] ACC sensöründen veri toplamış ve semantik insan FA'sını tanımak için iki yönlü yarı denetimli öğrenme temelli bir yöntem tasarlamışlardır. Elliott vd. [24], vücudun çeşitli kısımlarına bağlı 9 farklı ACC kullanarak SVM ve kNN algoritmaları ile % 97.6 doğruluk elde etmişlerdir. Jain ve Kanhangad [25] bir akıllı telefonda ACC ve Jiroskop (GYR - Gyroscope) sensör verilerini kullanmışlardır. Öznitelik seviyesi ve puan seviyesi füzyon tekniklerini kullanmışlardır. ACC verilerini kullanarak FA'yı tanımak için yöntemler geliştirilse de, literatür taramasına göre, bunların hiçbiri FA'nın ortamını tanımak için deneysel olarak uygulanmamıştır. Ayrıca, manyetik alanın yoğunluğu veya ortamın eğimi, içeride veya dışarıda olmasına bağlı olarak değişebileceği için, Magnetometre (MAG) ve Eğimölçer (CLI - Clinometre) sensörlerinin kullanılması FA'yı tanımak için ek bilgi sağlayabilir.

Son zamanlarda, çeşitli insan faaliyetlerini tanımak için birçok çalışma geliştirilmiştir. Bu çalışmalar FA'yı tanımak için giyilebilir cihazlar ve cep telefonları kullanmaktadır. Incel vd. [26] akıllı telefonları kullanarak FA tanıma yöntemini araştırmışlardır. Esfahani vd. [27], akıllı tekstil sensörlerini (STS) kullanarak yeni bir giyilebilir fanilin tasarımını ve geliştirilmesini tarif etmişlerdir. Mukhopadhyay [28], giyilebilir insan etkinliği algılama sistemlerinin genel performansını geliştirmek için sekiz bileşen düşünmüştür. Akıllı telefonlara yerleştirilmiş atalet sensörleri (örneğin ACC ve GYR), yürüme, koşma, oturma, ayakta durma, uzanma ve yukarı ve aşağı yürüme gibi FA'ları tanımak için kullanılır [29]. Farklı cihazlar, farklı yerleştirme sahaları ve farklı analiz türlerini kullanan çalışmaları karşılaştırmak oldukça güçtür [30].

Fiziksel aktivite tanıma çalışmalarının özet tablosu Çizelge 2.1'dedir.

Çizelge 2.1 Fiziksel Aktivite Tanıma Çalışmaları

Çalışma	Sensörler	Yöntemler	Veri Kümeleri
Kansız ve Diğ. [11]	Zaman tabanlı öznelikler	Naive Bayes, J48, K-star	Kendi veri kümeleri
Siirtola ve Röning [12]	İvme ölçer	Quadratic discriminant analiz	Kendi veri kümeleri
Seneviratne ve Diğ. [13]	İnceleme Çalışması		
Su ve Diğ. [14]	İnceleme Çalışması		
Bulling ve Diğ. [15]	İnceleme Çalışması		
Yang ve Diğ. [16]	İvme ölçer, jiroskop	Deep Convolutional Sinir Ağları	UCI HAR
He ve Diğ. [18]	İvme ölçer, jiroskop	CNN ve DeepConvLSTM	UCI HAR
Ordóñez ve Diğ. [19]	İvme ölçer, jiroskop, manyetometre	Deep Convolutional and LSTM Recurrent Neural Networks	OPPORTUNITY, PAMAP, Skoda and mHealth
Ha ve Diğ. [20]	İvme ölçer, jiroskop	Deep Convolutional Sinir Ağları	Mhealth
Wang ve Diğ. [22]	İvme ölçer	EEMD-Based Öznelikler ve Game-Theory-Based Öznelikler ile temel sınıflandırma yöntemleri	Kendi veri kümeleri
Lv ve Diğ. [23]	İvme ölçer	Semantik yarı denetimli sınıflandırma	Kendi veri kümeleri
Fullerton ve Diğ. [24]	Çoklu İvme ölçer	SVM	Kendi veri kümeleri
Jain ve Kanhangad [25]	İvme ölçer, jiroskop	SVM ve kNN	UCI HAR
Incel ve Diğ. [26]	İnceleme Çalışması		
Esfahani ve Nussbaum [27]	Tekstil algılayıcısı, ivme ölçer	Yapay sinir ağları	Kendi veri kümeleri
Mukhopadhyay [28]	İnceleme Çalışması		
Shoaib ve Diğ. [29]	İnceleme Çalışması		
Schrack v eDiğ. [30]	İnceleme Çalışması		

## 2.2. Uyku Apnesi ve Kalp Krizi Tanıma

Uyku Apnesi (OSA) uyku bozukluğu hastalığıdır. Bu hastalıkta uyku sırasında nefesin geçici olarak durması olur. Bu nefes durması genellikle 10 ila 30 sn

arasındadır. Bununla beraber, bu problem uyku sırasında defalarda olur [31]. Ağır bir şekilde olan bu problem, gündüz uyuklamasına, yorgunluğa ve azalmış hafızaya yol açabilir. OSA ayrıca anjina, aritmi ve inme ile ilgili olabilir. Birçok endüstriyel ve otomobil kazasında, birçok can kaybına ve milyarlarca dolara neden olan bir suçlu olarak kabul edilir.

Toplam ABD nüfusunun yaklaşık % 6,62'sine OUA teşhisi konmuştur [32]. OSA'nın ölüm oranını arttırdığı, ayrıca yaşam kalitesini ve üretkenliği azalttığı bilinmektedir. OUA yaşayan kişilerin hipertansiyon, diyabet, depresyon, obezite ve kanser gibi kronik hastalıklardan muzdarip olma olasılığı daha yüksektir. ABD'de 18 milyon insanın OSA'sı olduğu, OSA'lı olarak yaklaşık 63 milyon kişinin yaşadığı ve henüz klinik olarak teşhis edilmediği tahmin edilmektedir. OSA izleme ve tespit araştırması literatürde büyük miktarda dikkat çekmiştir. OSA esnasında nabız, solunum hızı ve oksijen satürasyon (SpO<sub>2</sub>) verileri ile oluşturulan modeller ile OSA saptama yaklaşımları denenmiştir [33-35].

Bu hastalığı teşhis etmek için kullanılan birincil method Polisomnografi (PSG) testine tabi tutulmaktır. PSG, vücuda bağlı ve bir uyku laboratuvarında gerçekleştirilen en az on bir sensör ile (yirmi iki tel) kapsamlı bir testtir. Apne Hipopne İndeksi (AHI) saatte beşten fazla ise hastanın OSA olduğu söylenir. AHI'nin tanımı, kan oksijeninde desatürasyon ile belirli bir süre solunum durur. Bu desatürasyon ve solunum problemi ile hem uyku bozukluklarını hem de yetersiz oksijen şiddeti değerlendirilebilir. Ağırlık kaybı, sürekli pozitif hava yolu basıncı (CPAP) [36] ve üst hava yolu ameliyatı dahil olmak üzere Uyku Apnesi için çeşitli tedaviler vardır. Polisomnografi (PSG) uyku apnesi de dahil olmak üzere uyku bozukluklarının teşhisinde standart hale gelmiştir. PSG'de nefes hava akımı, solunum hareketi, oksijen satürasyonu, elektroensefalogram (EEG), elektromiyogram (EMG), elektrookülogram (EOG) ve elektrokardiyogram (EKG) kaydı ile vücut pozisyonu bulunur. Ancak, laboratuvar ortamında doktorlar ve hemşirelerin gözetimi altında tam gece uykusu ile PSG yapılır [37]. Her ne kadar PSG OUA teşhisinde çok önemli olarak düşünülse de, PSG bazı araştırmacılar tarafından rahatsız edici, pahalı ve sınırlı kullanılabilirlik gibi birçok eleştiri almıştır [38]. Bu nedenle, PSG'nin yapabildiklerini yapan ve hastanede olan bir kurulum gerektirmeyen kolay yöntemlere ihtiyaç vardır.

EKG kaydı, uyku bozuklukları tespitinde en basit ve etkili teknolojidendir. EKG sinyallerinin kalp atış hızı aralıklarını (RR) yenmek için atımdaki sıklık varyasyonlarının, apne sırasında bradikardi ve bunu takiben taşikardi ile oluşan uyku apnesi olayları ile ilişkili olduğu bildirilmiştir. Bu tasarım klinik uyku apnesi belirtileri olan hastaları saptamada başarılı bir şekilde kullanılmıştır [34]. Çeşitli çalışmalar, birkaç yeni yöntemin, uyku apnesini kalp hızı değişkenliği (HRV) değişikliklerinden tanınabileceğini doğrulamıştır [39]. Chazal tarafından, RR-aralıklarından ve EKG'den türetilmiş solunum sinyalinden (EDR) elde edilen değişkenlerin istatistiksel ölçümünü kullanarak obstrüktif uyku apnesi yöntemi önerilmiştir [38]. 2010 yılında, tek bir EKG kullanarak uyku aşamalarında sınıflandırma ve apne araştırmaları yapan Yılmaz, RR-aralığından, yani medyan, çeyrekler arası aralıktan (IQR) ve ortalama mutlak sapmadan (MAD) türetilmiş üç öznitelik kullanmıştır. Sınıflandırma, kNN, Quadratic Discriminant Analysis (QDA) ve Support Vector Machines (SVM) kullanılarak yapılmıştır [40].

Günümüzde uygulanan OSA algılama yöntemlerinin çoğu, zaman bölgesi, frekans alanı ve diğer doğrusal olmayan özniteliklerin çeşitli EKG-bazlı sinyallerden çıkarılmasına ve daha sonra OSA oluşumunu belirlemek için bu özelliklere sahip sınıflandırıcıların oluşturulmasına odaklanmaktadır. Bir kişinin uykusu sırasında tüm EKG'sini incelemek popüler bir araştırma konusudur. Örneğin, Khandoker vd., OSR hasta tanımlaması için, RR ters çevirmelerinden (EKG'lerden elde edilen R dalgası aralıkları olan) ve EKG'den türetilmiş solunum (EDR) sinyallerinden dalgacık ayrışımı ile çıkarılan çeşitli özniteliklere dayanan bir SVM önermiştir. Çalışmalarında, test kümesindeki deneklerin % 90'ından fazlası doğru bir şekilde sınıflandırılmıştır [41].

OSA tanısı için her bir eşit uzunluktaki EKG sinyali segmentinden öznitelikler çıkaran birkaç alternatif sınıflandırma yöntemi geliştirilmiştir. Bu yöntemler, bir hastanın genel OSA şiddetinin yanı sıra her segmentte OSA oluşumunu belirleyebilir. Shouldice vd.'de, pediatrik deneklerde OSA'yı saptamak için dakikalara bölünmüş EKG sinyallerini kuadratik diskriminant analizine (QDA) entegre ederek bir sınıflandırma yöntemi önermişlerdir [42]. Böylelikle bölüm başına bazda % 72.1 ve % 84 oranında bir doğruluk elde edilmiştir. Mendex vd., her segmentten 70'in üzerinde öznitelik üzerinde çalışmış ve en yakın komşu (kNN) ve sinir ağları ile sınıflandırma doğruluğunu % 88 bulmuştur[43]. Bsoul vd., akıllı

telefonlara dayalı gerçek zamanlı bir apne izleme sistemi geliştirmiştir [44]. RR aralıklarından 63 özellik ve EDR sinyallerinden 48 dakika çıkarmışlardır ve apne tespiti için SVM'ler kullanmışlardır.

Birçok OSA tespit çalışmasındaki iki yaygın sorun, yüksek boyutsal öznitelik alanı ve karar verme sürecindeki zorluklardır. Literatürde, öznitelik aralığının boyutunu azaltmak için istatistiksel değerlendirme [44], sarmalayıcı yöntemleri [43] ve temel bileşen analizi dahil olmak üzere çeşitli öznitelik seçim yöntemleri uygulanmıştır [45].

EKG, EEG, Oksimetre, burun hava akımı ve benzeri gibi biyo-sinyal sensörünü kullanarak uyku apnesi teşhisini kolay ve rahat hale getirmek için önemli bir araştırma girişimi Oliver ve Flores- Mangas tarafından yapılmıştır [46]. OSA taraması için kan oksijen seviyesini ( $SpO_2$ ) kullanmışlardır. Fu-Chung vd. [47], OSA tanımlamak için entegre CPAP hava akımı sinyali kullanmışlardır. Pentagay vd. [48], EKG ile birlikte OSA bölümleri sırasında üretilen kalp sesini (S1) kullanmışlardır. Patangay vd. ile Raymond vd. SVM sınıflandırıcılarının apne taramasında uygulanması konularını tartışılmaktadır [49, 50]. EKG'ye dayanan diğer OSA tarama çalışmaları, Gauss sınıflandırıcılarını, doğrusal ya da ikinci dereceden ayırmacılarını kullanmışlardır [51, 52].

Sani vd. insan koluna sensör takıp kandaki Critine Kinase, Myoglobin, Troponin I, Troponin T değerlerini toplamışlardır [53]. Buradaki bilgiler ile eşik değerlerini geçer geçmez cep telefonu ile haberleşip sunucuya yazmışlardır. Eşik değerlerini geçince kalp krizi esnasında olduğu bilgisi bulunuyor ve daha sonra uzaktaki sistemler ile alarm vermiştir. Ravish vd. EKG verisinden kalp krizi olma durumunu tahmin etmişlerdir [54]. Yapay sinir ağları ve genetik algoritma kullanarak tahmin etmişlerdir. EKG verisi yanında kolestrol, tansiyon gibi verileri de hesaplamaya dahil etmişlerdir. Uyku apnesi tanıma çalışmalarının özet tablosu Çizelge 2,2'de gösterilmiştir.

Literatür göz önüne alındığında, cep telefonları, vücut sensörlerinin kullanımı ve akıllı telefonlar yaygınlaştıktan sonra insan faaliyetlerini sınıflandırmak için kullanılmıştır. Bu çalışmalar genellikle ACC ve GYR sensörlerini kullanır. Bunlara ek olarak, bu çalışmada MAG, CLI ve bir oksimetre sensörü kullanılmıştır. Cihaz çeşitliliğindeki zorlukları görerek metodolojiyi literatürdeki veri kümeleri ile de

karşılaştırılmıştır. Bununla birlikte, bilinen kadarıyla, bir FA'nın açık ya da kapalı alan bilgisi, literatürdeki hiçbir çalışmada özellikle yaşlıların sağlığı için çok önemli olabilecek SpO<sub>2</sub> sensörü kullanılarak tahmin edilmemiştir. Bu çalışmanın amacı, insan FA'sının açık ve kapalı alan bilgisini, mobil telefon ve oksijen satürasyon sensörü kullanarak tespit edilmesidir. Ayrıca, çoklu sensör kullanımının, açık ve kapalı alan ile birlikte FA tespitinin doğruluğu üzerindeki etkisini araştırılmıştır.

Çizelge 2.2 Uyku Apnesi Tanıma Çalışmaları

Çalışma	Sensörler	Yöntemler	Veri Kümeleri
Penzel ve Diğ. [33]	EKG (nabız, nabız enerjisi, S amplitude)	Hidden Markov Model	Kendi veri kümeleri
Mendez ve Diğ. [35]	EKG	Wavelet Analizi	Kendi veri kümeleri
Yılmaz ve Diğ. [40]	Polisomnograf	quadratic discriminant analysis ve SVM	Kendi veri kümeleri
Khandoker ve Diğ. [41]	EKG ve Oximetre	SVM	Sleep Research Unit (SRU) Database ve Physionet Apnea-EKG Database
Shouldice ve Diğ. [42]	EKG	quadratic discriminant analysis	Kendi veri kümeleri
Mendez ve Diğ. [43]	EKG	kNN ve Yapay Sinir Ağı	Kendi veri kümeleri
Bsoul ve Diğ. [44]	EKG	SVM	Kendi veri kümeleri
Isa ve Diğ. [45]	EKG	kNN, Naive Bayes, Linear SVM, RBF SVM	Physionet Apnea-EKG Database
Fu-Chung ve Diğ. [47]	Konuşma ve ses verisinden	probability density functions	Kendi veri kümeleri
Khandoker ve Diğ. [48]	EKG	SVM	Kendi veri kümeleri
Patangay ve Diğ. [49]	Kalp sesi ve EKG	SVM	Kendi veri kümeleri
Raymond ve Diğ. [50]	EKG ve Oximetre	SVM	Physionet Apnea-EKG Database
Chazal ve Diğ. [51]	EKG	quadratic discriminant analysis	Physionet Apnea-EKG Database ve Philipps-University Database

### 3. KULLANILAN TEMEL YÖNTEMLER

#### 3.1. Sınıflandırma Algoritmaları

##### 3.1.1. K En Yakın Komşu (kNN) Algoritması

Kullanım basitliği ile çok sık kullanılan sınıflandırma algoritmaları içerisinde bulunan K en yakın komşu algoritmasında, veri kümesindeki elemanların birbirlerine olan mesafelerinin bulunması ile sınıflandırma işlemleri yapılmaktadır [69]. En yakın komşular (kNN) algoritması, hem sınıflandırma hem de regresyon sorunlarını çözmek için kullanılabilir basit, uygulaması kolay, denetimli bir makine öğrenme algoritmasıdır. kNN, fonksiyonun sadece yerel olarak yaklaştığı ve tüm hesaplamaların sınıflandırmaya kadar ertelendiği, bir temele dayalı öğrenme veya tembel öğrenmedir. kNN algoritması, tüm makine öğrenmesi algoritmalarının en basitleri arasındadır.

Elimizdeki eğitim kümesine göre sınıflandırılacak olan test verisinde, sınıflandırma işlemi K değeri ve uzaklık algoritmasından gelen sonuca göre değişmektedir. Bir sınıflandırma probleminde, sınıfı bilinmeyen bir öge, komşularının çoğunluk oyuyla sınıflandırılır; test verisi, uzaklık fonksiyonu ile ölçülen en yakın komşuları arasında en yaygın olan sınıfa atanır. K = 1 ise, test verisi en yakın komşusunun değerini alır. K en yakın komşu algoritmasında yaygın olarak farklı uzaklık algoritmaları kullanılmaktadır. Bunlar Öklid Uzaklığı, Manhattan vb. uzaklığıdır.

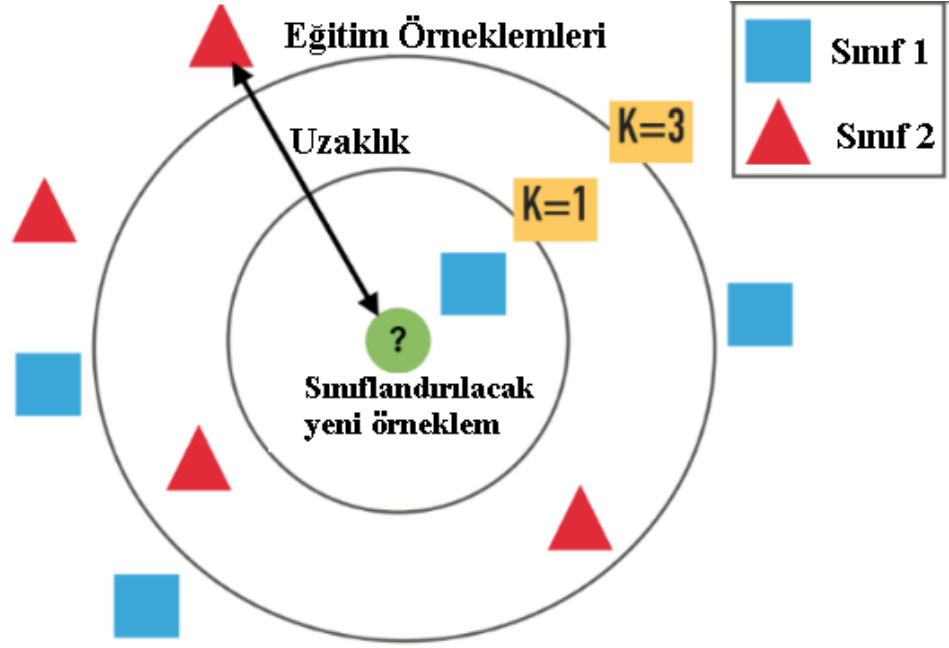
Matematikte, Öklid mesafesi veya Öklid metriği, birinin bir cetvelle ölçeceği iki nokta arasındaki "sıradan" mesafedir ve Pisagor formülüyle verilir. Bu formülü mesafe olarak kullanarak, Öklid uzayı bir metrik boşluğa dönüşür. İlişkili norm, Öklid normu olarak adlandırılır. n boyutlu Öklid Uzayında  $A = (a_1, a_2, \dots, a_n)$  ve  $B = (b_1, b_2, \dots, b_n)$  noktaları arasındaki Öklid uzaklığı (3.1) eşitliği ile hesaplanır.

$$\sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_n - b_n)^2} = \sqrt{\sum_{i=1}^n (a_i - b_i)^2} \quad (3.1)$$

Manhattan Uzaklığında iki nokta arasındaki mesafe, eksenler boyunca dik açılarda ölçülmektedir. İki nokta arasındaki Manhattan mesafesi (3.2) eşitliği ile hesaplanır.

$$\sum_{i=1}^n |a_i - b_i| \quad (3.2)$$

Bu çalışmada deneyler yapılırken basitliği düşünülerek Öklid Uzaklığı kullanılmıştır. k en yakın komşular algoritmasının kullanımına örnek Şekil 3.1'de verilmiştir.



Şekil 3.1 k En Yakın Methodu Örneği

Yukarıdaki şekli incelediğimizde, test numunesi (daire içinde) birinci sınıf karelere veya ikinci sınıf üçgenlere sınıflandırılmalıdır. Eğer  $k = 3$  ise (dış daire) ikinci sınıfa atanır, çünkü iç çember içerisinde 2 üçgen ve sadece 1 kare vardır. Örneğin,  $k = 5$  ise birinci sınıfa atanır (dış çemberin dışındaki 3 kare ve 3 üçgen).

Öklid Uzaklığı kullandığımız çalışmamızda,  $k = 1, 3, 5$  ve  $7$  deneylerine dayanarak sınıflandırma sırasında  $k = 3$  seçilmiştir.

### 3.1.2. Naive (Saf) Bayes Sınıflandırma Algoritması

Naive Bayes sınıflandırması benzer şekilde basitliği ile sık kullanılan bir algoritmadır. Bu algoritma istatistiksel methodları kullanır. Bayes Teoremini [69] bu istatistiksel yöntemler içerisinde. Naive Bayes, sınıflandırıcı oluşturmak için basit bir tekniktir. Sınıf etiketlerini bazı sonlu kümelerden çizilen öznitelik değerlerinin vektörleri olarak temsil edilen, sorunlu durumlara sınıf etiketlerini atayan bir modeldir. Bu sınıflandırıcıları eğitmek için tek bir algoritma yoktur, ancak ortak bir ilkeye dayanan bir algoritma ailesi vardır. Tüm Bayes sınıflandırıcıları, belirli bir özniteliğin değerinin, sınıf değişkeni verilen herhangi bir özniteliğin değerinden bağımsız olduğunu varsayar. Örneğin, bir meyve kırmızı, yuvarlak ve çapı yaklaşık 10 cm ise elma olarak kabul edilebilir. Saf bir Bayes sınıflandırıcısı, bu özniteliklerin her birinin, renk, yuvarlaklık ve çap özellikleri arasındaki muhtemel korelasyonlara



bakılmaksızın, bu meyvenin bir elma olma ihtimaline bağımsız olarak katkı sağladığını düşünmektedir.

Bazı olasılık modelleri için Naive Bayes sınıflandırıcıları denetimli bir öğrenme ortamında çok verimli bir şekilde eğitilebilir. Birçok pratik uygulamada, saf Bayes modelleri için parametre tahmini, maksimum olasılık yöntemini kullanır. Başka bir deyişle, Bayes olasılığını kabul etmeden veya herhangi bir Bayes yöntemini kullanmadan saf Bayes modeliyle çalışılabilir.

Bayes yönteminin temeli olan Bayes teoremi, olasılık kuramı içinde incelenen önemli bir konudur. Bu teorem bir rastgele değişken için olasılık dağılımı içinde koşullu olasılıklar ile marjinal olasılıklar arasındaki ilişkiyi gösterir. Bayes teoreminin formülü (3.3) eşitliğinde verilmiştir.

$$P(A_j|B) = \frac{P(A_j \cap B)}{\sum_{i=1}^k P(A_i \cap B)} = \frac{P(A_j) P(B|A_j)}{\sum_{i=1}^k P(A_i) \cap P(B|A_i)} \quad (3.3)$$

İlgili eşitlikte  $A_1, A_2, \dots, A_k$  olayları ayrık ve  $S$  örnek uzayını bütüne tamamlayan olaylar olsun.  $P(A_i) \neq 0, i = 1, 2, \dots, k$  ve  $B$  örnek uzay  $S$ 'den bir olay şeklindedir.

Bu çalışmada, Matlab Statistics Tools ve makine öğrenmesi araçlarını kullanarak Gauss dağılımına dayanan Naive Bayes algoritması uygulanmıştır. Bu yöntemde, ortalama ve kovaryans matrisi, eğitim verileri kullanılarak hesaplanır ve test aşamasında, kategorilerin koşullu olasılıkları verilen bir model için hesaplanır. Daha sonra, sonsal olasılık (posterior probability) hesaplanır. Test deseni, sonsal olasılığın en yüksek olduğu sonsal olasılık kullanılarak belirli bir kategoride sınıflandırılır.

### 3.1.3. Destek Vektör Makineleri (SVM)

Destek vektör makineleri algoritmasında eğitim veri kümesine göre, test edilecek olan verilerde sınıfların birbiri ile tam olarak ayırmasını sağlayan yapının oluşturulması hedeflenmektedir [70]. Eğer veri kümesinde iki sınıf varsa, verilerin birbiri ile ayırma işlemi (sınıflandırma işlemi) bir doğru yardımı ile olacaktır. Eğer veri kümemizde birden fazla sınıf bulunuyorsa verileri birbiri ile ayırma işlemi (sınıflandırma işlemi) doğrusal olmayan çizgilerle ayrıştırılır.

SVM, verileri doğrusal bir hiper düzlemlerle iki kategoriye ayıran bir sınıflandırıcıdır. SVM, eğitim verilerini alır ve bir kategorinin parçası olarak işaretler ve test örneğinin

mevcut bir sınıfın üyesi olup olmadığını tahmin eder. SVM modelleri, verileri bir hiper düzleme bölünmüş alandaki bir nokta olarak gösterir. SVM'nin en önemli avantajları, kullanımı basit ve yüksek hassasiyetlidir. Çekirdek parametrelerini optimize etmek için SVM çekirdeği ve grid arama algoritması olarak radyal temel işlevi (RBF) kullanılmıştır.

SVM algoritmasının uygulanması için Libsvm Kütüphanesi kullanılmıştır [71]. SVM eğitimi bir veya daha fazla parametre ile kontrol edilir. Bazen gevşeklik parametresi olarak adlandırılan C parametresi, eğitim numunelerinin kaç tanesinin karar sınırının yanlış tarafında bulunmasına izin verildiğini kontrol eder. Bazı eğitim örneklerinin sınırdan sızmasına izin vermek verilere uymanın azaltılmasına yardımcı olabilir. Eğitim modeli, eğitim örneklerinin %100'ünü sınıflandıran bir sınır seçtiğinde aşırı yerleştirme (overfit) oluşabilir. Overfit olma durumunda, model iyi bir sınıflandırıcı olmayacaktır, çünkü sınıfları o kadar dar bir şekilde tanımlamaktadır ki, sınıfları eğitim vektörleriyle tam olarak eşleşmeyen veriler için iyi bir tahmin yapamamıştır. Overfit durumundan kaçınarak, durgunluk (slack) parametresi %100 altındaki eğitim doğruluğu ile tahmin doğruluğu arasında bir değişim sağlar.

Bu nedenle, C'nin yanı sıra diğer parametrelerin en iyi (optimal) değeri araştırılmıştır. Slack parametresine ek olarak, çeşitli çekirdek fonksiyonlarıyla ilişkili parametreler vardır. Bu çalışmada, ikinci bir parametrenin  $\gamma$  belirtilmesini gerektiren Radyal Temel İşlev (RBF) çekirdeği kullanılmıştır. (C,  $\gamma$ ) için en uygun değerleri bulmak için, kullanılan veri kümesi üzerinde bir en iyi parametre optimizasyonu araştırması yapılmıştır.

Hiper düzlemin doğrusal SVM'de öğrenilmesi, problemin bir miktar lineer cebir kullanılarak dönüştürülmesiyle yapılır. Çekirdeğin rolü olduğu yer burasıdır. Doğrusal çekirdek için, giriş ( $x$ ) ve her bir destek vektörü ( $x_i$ ) arasındaki nokta çarpımını kullanan yeni bir girişin tahmin denklemi (3.4) eşitliğinde gösterilmiştir.

$$w_0 \cdot x + b_0 = 0 \quad (3.4)$$

(3.4) eşitliği içerisinde  $w$  ile vektör  $b$  ile sabit değer ifade edilmektedir. Kullandığımız veri kümesinde, sınıflandırma aşamasında düzlemsel olarak sınıflandırma işlemi yapamadığımız durumlarda çekirdek (kernel) fonksiyonlar kullanılmıştır. Çekirdek fonksiyon kullanıldığı durumlarda veri kümesinde birden fazla boyutlu bir uzay temsili sağlanarak, ayırt edici düzlemler ortaya çıkarılır. Bu sayede veri kümesinin

dağılımı birden fazla boyuttaki uzayda görüntülenebilmektedir. Çekirdek fonksiyonlarının tanımlanması destek vektör makineleri için eşitlik (3.5)'deki gibidir.

$$K(x_i, x_j) = \varphi(x) \cdot \varphi(x_j) \quad (3.5)$$

(3.5) eşitliği kullanıldığında eğer veri kümesi doğrusal olarak ayrılamıyorsa, kullanılan veri kümesi için sınıflandırma kararı (3.6) eşitliği ile bulunan fonksiyon kullanılarak yapılır.

$$f(x) = \text{sign}(\sum_{i=1}^n y_i \cdot \alpha_i \cdot \varphi(x) \cdot \varphi(x_i) + b) \quad (3.6)$$

Destek vektör makinelerinde kullanılan çekirdek fonksiyonlarına bakıldığında,

- $d$  dereceli Polinom çekirdek fonksiyonu (3.7) eşitliği ile,

$$K(x, x_i) = (\gamma \cdot x_i + r)^d \quad (3.7)$$

- Sigmoid çekirdek fonksiyonu (3.8) eşitliği ile,

$$K(x, x_i) = \frac{(\gamma \cdot x_i + r)^d}{\sqrt{(\gamma \cdot x_i + r)^d (\gamma \cdot x_i + r)^d}} \quad (3.8)$$

- $\gamma$  çekirdek boyutuna sahip radyal tabanlı çekirdek fonksiyonu (3.9) eşitliği ile gösterilmiştir.

$$K(x, x_i) = e^{-\gamma \|x - x_i\|^2} \quad (3.9)$$

Burada  $r$ ,  $d$ , ve  $\gamma$  işaretleri kernel parametrelerini göstermektedir.

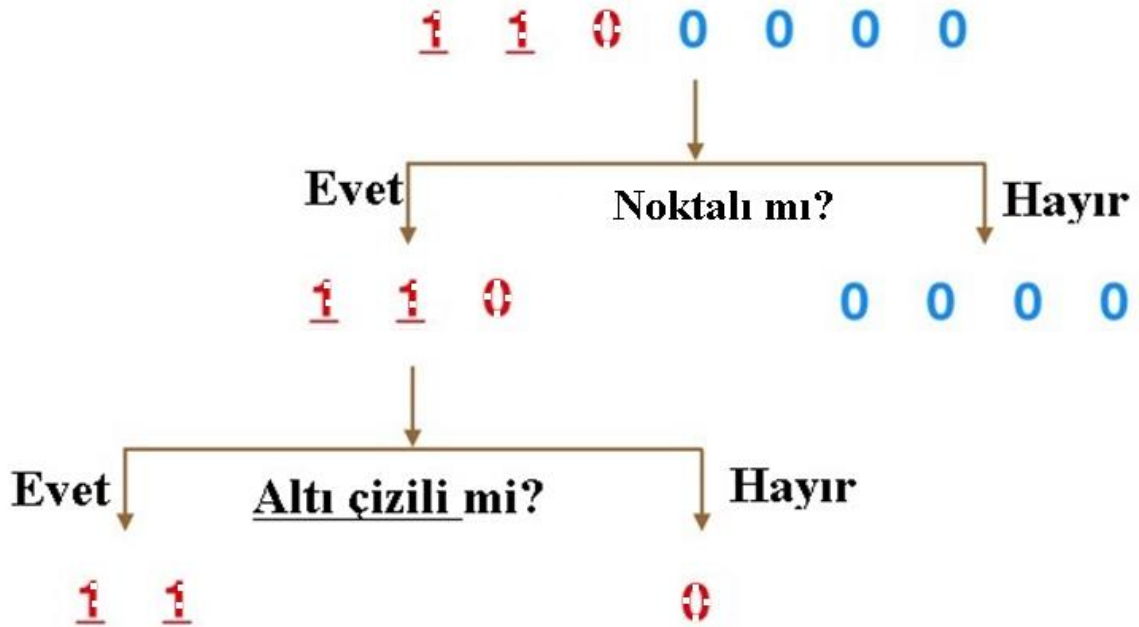
#### 3.1.4. Rastgele Orman (Random Forest)

Rastgele orman (Random Forest), başarılı sonuçlar vermesinden dolayı çok sık kullanılan makine öğrenmesi algoritmaları arasındadır. Sınıflandırma dışında regresyon problemlerine uygulanabilir olmasından dolayı da kullanım alanı geniştir [72]. Kullanıcının, tek sınıflandırma ağaçlarının yapısını bildiğini varsaydığımızda Rastgele Ormanlarda birçok sınıflandırma ağacı bulunur. Bir giriş vektöründen yeni bir nesneyi sınıflandırmak için, giriş vektörünü ormandaki ağaçların her birine yerleştiririz. Her ağaç bir sınıflandırma sonucu verir ve ağacın o sınıfa "oy verdiğini" söyleriz. Orman, ormandaki tüm ağaçların üzerinde en fazla oyu alan sınıflandırmayı seçer. Her ağaç aşağıdaki gibi yetiştirilir:

- Eğitim kümesindeki vakaların sayısı  $N$  ise, örnek vaka rastgele ve orijinal verilerden seçilir. Bu örnek ağacın yetiştirilmesi için hazırlanmış bir eğitim olacaktır.
- $M$  giriş değişkenleri varsa, her düğümde,  $M$  değişkenlerinin  $M$ den rastgele seçildiği ve bu  $m$  üzerindeki en iyi bölmenin düğümü bölmek için kullanıldığı şekilde  $m \ll M$  sayısı belirtilir. Ormancılıkta  $m$  değeri sabit tutulur.
- Her ağaç budama yapılmadan mümkün olan en büyük ölçüde yetiştirilir.

Karar ağaçları sınıflandırma algoritmalarının problemlerinden biri aşırı öğrenmedir (overfitting). Rastgele orman bu problemi aşmak amacıyla veri kümesinden ve öznelik kümesinden rastgele olarak farklı sayılarda alt kümeler seçerek modeli eğitir. Bu yöntemle farklı sayılarda karar ağacı oluşturur ve bu ağaçlar kendilerince tahminde bulunur. Sonuç olarak eğer regresyon ile tahminde bulunuyorsak ağaç tahminlerini ortalama olarak kullanırız, sınıflandırma için tahmin yapıyorsa ağaç tahminlerinden en çok oy alan seçilir.

Rastgele orman algoritmasının temelinde karar ağaçları bulunmaktadır. Bir karar ağacının bir örnek üzerinde aşağıdaki şekildeki gibi inceleyebiliriz (Şekil 3.2).



Şekil 3.2 Karar ağacı örneği

Veri kümesinin, Şekli 3.2'nin üstündeki sayılardan oluştuğunu düşünelim. İki tane 1 ve beş tane 0'ımız vardır (1 ve 0 bizim sınıfımız) ve sınıflarını özelliklerini kullanarak ayırma isteğindeyiz. Öznitelikler noktalı ya da değildir ve bazılarının altı çizilidir.

0'lardan biri dışında hepsi noktasız olduğu için ayrılmak için ilk özellik olarak seçebiliriz. Böylece ilk düğümümüzü bölmek için "noktalı mı?" sorusunu kullanabiliriz. Ağaçtaki bir düğümü, yolun ikiye ayrıldığı kısım olarak düşünebiliriz. Noktalı olanlar dalından aşağı inebilir durumdadır ve noktasız olanların hepsinin sınıfı aynı olduğu için ağaç orada sonlanmıştır. Daha sonra da altı çizili mi sorusunu sorulduğunda sınıfların tamamıyla ayrılmış olduğu görülmektedir.

Rastgele orman da, topluluk olarak çalışan çok sayıda bireysel karar ağacından oluşur. Rastgele ormandaki her bir ağaç, bir sınıf tahminde bulunmakta ve en fazla oy alan sınıf, modelimizin öngörüsü haline gelmektedir.

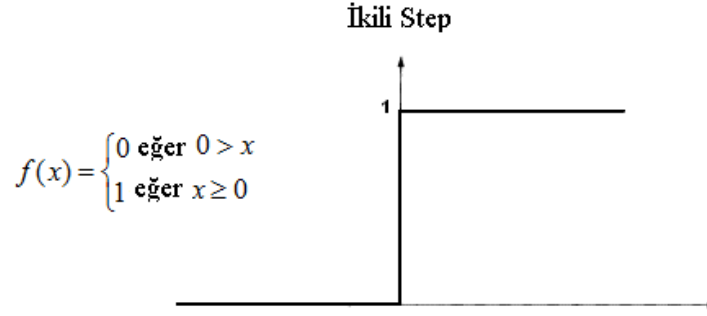
### **3.1.5. Yapay Sinir Ağları (YSA)**

Yapay Sinir Ağları, insan beynindeki sinir yapılarından esinlenerek modellenen, kalıpları tanımak için tasarlanmış bir algoritmalar kümesidir. Verileri kullanırken bir tür makine öğrenmesi algoritması olarak, sınıflandırma veya kümeleme yoluyla yorumlarlar. Yapay sinir ağları kümelemede ve sınıflandırmada kullanılabilir. Bunlar, verilerin üstünde bir kümeleme veya sınıflandırma katmanı olarak düşünülebilir [69]. Eğer etiketlenmemiş veriler varsa, örnek girdiler arasındaki benzerliklere göre kümelemeye yardımcı olurlar. Üzerinde çalışacak etiketli bir veri kümesi olduğunda verileri sınıflandırırlar.

YSA'larda sinyaller (veriler) arasındaki iletimler nöronlar vasıtasıyla olur. Nöronlar her sinyale tepki vermemektedir. Tepki vermek için eşik değerleri bulunmaktadır. Bu eşik değerlerini kullanarak çıktı üreten fonksiyonlara Aktivasyon Fonksiyonları denir.

#### **3.1.5.1. Adım aktivasyon fonksiyonu**

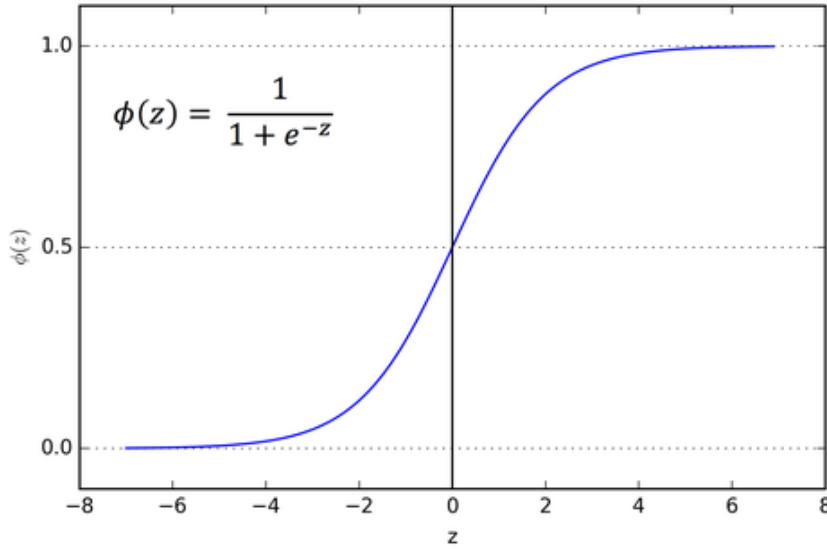
İkili adım aktivasyon fonksiyonunda girdinin 0'dan büyük ve küçük olma durumuna göre çıkış değeri ayarlanmaktadır. Şekil 3.3'te ikili adım fonksiyonu ve grafiği görülmektedir.



Şekil 3.3 İkili adım aktivasyon fonksiyonu gösterimi

### 3.1.5.2. Sigmoid aktivasyon fonksiyonu

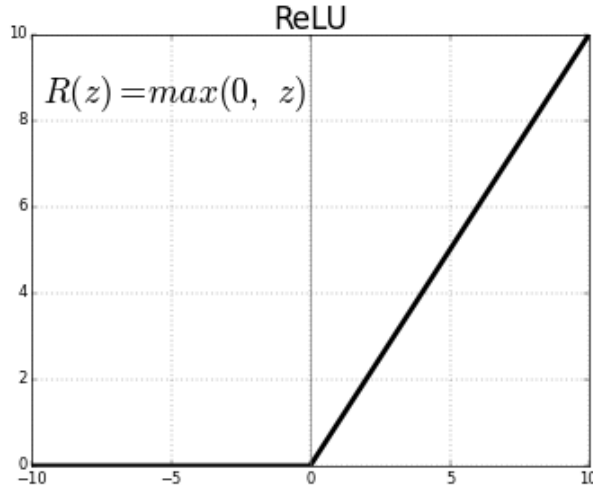
Sigmoid aktivasyon fonksiyonunu kullanmamızın ana nedeni, (0 - 1) arasında olmasıdır. Bu nedenle, özellikle bir olasılığı çıktı olarak tahmin etmemiz gereken modeller için sigmoid aktivasyonunu kullanmak doğru seçimdir. Şekil 3.4 sigmoid fonksiyonunun grafiğini göstermektedir.



Şekil 3.4 Sigmoid aktivasyon fonksiyonu gösterimi

### 3.1.5.3. Relu aktivasyon fonksiyonu

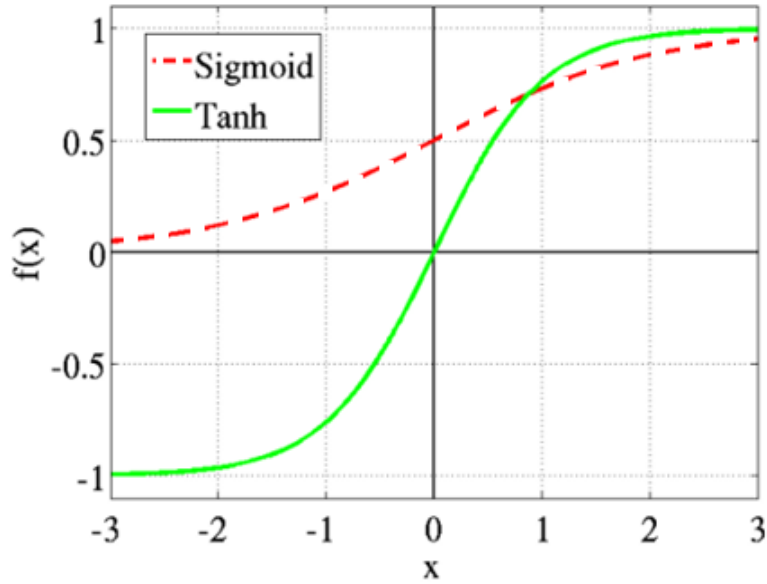
YSA modelleri geliştirilirken en çok kullanılan aktivasyon fonksiyonudur. Şekil 3.5 'de görüldüğü üzere sıfırdan küçük değerleri sıfır olarak değiştirmektedir.



Şekil 3.5 Relu aktivasyon fonksiyonu gösterimi

#### 3.1.5.4. Tanh aktivasyon fonksiyonu

Tanh fonksiyonunun aralığı -1 ile 1 arasındadır. Tanh fonksiyonunun avantajı, negatif girişlerin negatif eşleneceği ve sıfır girişlerinin tanh grafiğinde sıfıra yakın eşleştirileceğidir. Şekil 3.6'da Tanh ile Sigmoid arasındaki fark görülmektedir.



Şekil 3.6 Tanh ile Sigmoid aktivasyon fonksiyonları arasındaki fark

### **3.1.6. Çok Katmanlı Perceptron (Multilayer Perceptiron (MLP))**

Çok katmanlı perceptron (MLP), bir dizi girdiden bir çıktı kümesi üreten, ileriye dönük bir yapay sinir ağıdır [69]. Bir MLP, giriş ve çıkış katmanları arasında yönlendirilmiş bağlı birkaç giriş düğümü katmanı ile karakterize edilir. MLP, ağı eğitmek için geri yayılma yöntemini kullanır. MLP, çoklu katmanları bağlayan bir sinir ağıdır. Bu, düğümler boyunca sinyal yolunun sadece bir yöne gittiği anlamına gelir. Giriş düğümleri dışındaki her düğüm, doğrusal olmayan bir aktivasyon işlevine sahiptir. Bir MLP denetimli bir öğrenme tekniği olarak geri yayılımı kullanır. Birden fazla nöron tabakası olduğu için, MLP derin bir öğrenme tekniğidir. MLP, denetimli öğrenmenin yanı sıra hesaplamalı sinirbilim ve paralel dağıtılmış işlem araştırması gerektiren sorunları çözmek için yaygın olarak kullanılır.

### **3.1.7. Evrışimsel Sinir Ağı (ConvNet / CNN)**

Bir Evrışimsel Sinir Ağı (ConvNet / CNN), bir girdi alabilen, verideki çeşitli yönler / nesnelere önem veren (öğrenilebilir ağırlıklar ve önyargılar) ve birini birbirinden ayırabilen Derin Öğrenme algoritmasıdır [76]. CNN'de gereken ön işleme, diğer sınıflandırma algoritmalarına kıyasla çok daha düşüktür. İlkel yöntemlerde, filtreler el yapımı olsa da, yeterli eğitim ile CNN bu filtreleri / özellikleri öğrenebilir.

Bir CNN'in mimarisi, insan beynindeki nöronların bağlantı modelininkine benzer ve Görsel Cortex'in organizasyonundan ilham almıştır. Bireysel nöronlar, uyarıcılara sadece Reseptif Alan olarak bilinen görsel alanın sınırlı bir bölgesinde yanıt verir. Bu tür alanlardan oluşan bir koleksiyon, tüm görsel alanı kapsayacak şekilde üst üste biner.

CNN, sinyallerden özellikler çıkarma konusunda yetkindir ve görüntü sınıflandırma, konuşma tanıma ve metin analizinde umut verici sonuçlar elde etmiştir. Fiziksel aktivite tahmini gibi zaman serileri sınıflamasına uygulandığında, CNN'nin diğer modellere göre iki avantajı vardır: yerel bağımlılık ve ölçek değişmezliği. Yerel bağımlılık, fiziksel aktivite tanıma işlemindeki sinyallerin ilişkili olması muhtemeldir, ölçek değişmezliği farklı hızlar veya frekanslar için ölçek değişmezini ifade eder.

CNN'yi FA tahminine uygularken, göz önünde bulundurulması gereken birkaç husus vardır. Bunlar girdi adaptasyonu, havuzlama ve ağırlık paylaşımı (input adaptation, pooling, and weight-sharing)'dir.



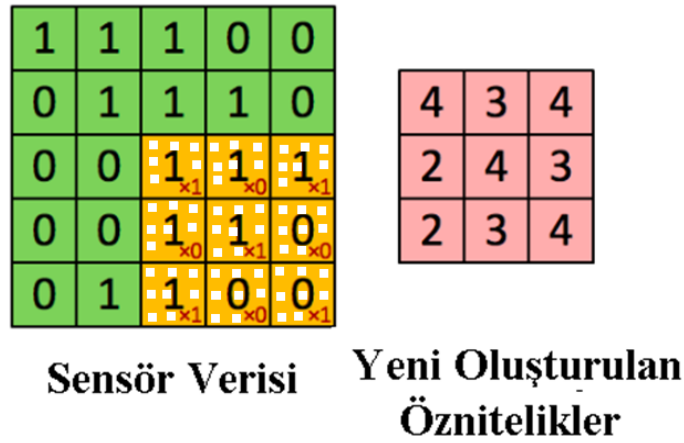
1) Giriş adaptasyonu: Görüntülerin aksine, çoğu fiziksel aktivite tahmininde kullanılan sensörler zamansal çok boyutlu okumaları olan ACC gibi zaman serisi okumaları üretir. Bu girişlere CNN uygulanmadan önce giriş uyarlaması gereklidir. Ana fikir, sanal bir görüntü oluşturmak için girdileri uyarlamaktır.

2) Havuzlama: Evrişim havuzu ( convolution-pooling ) birleşimi CNN'de yaygındır ve çoğu yaklaşım evrişimden sonra maksimum ya da ortalama havuzda gerçekleştirdi [77]. Aşırı donatmadan kaçınmanın yanı sıra, havuzlama aynı zamanda büyük verilerle ilgili eğitim sürecini de hızlandırabilir.

3) Ağırlık paylaşımı: Ağırlık paylaşımı [36], yeni bir görevdeki eğitim sürecini hızlandırmak için etkili bir yöntemdir.

Konvolüsyonlu Sinir Ağı (ConvNets veya CNN) mimarisinde üç önemli fikir vardır: seyrek etkileşimler, parametre paylaşımı ve eşdeğer temsiller [76]. Konvolüsyondan sonra, sınıflandırma veya regresyon görevlerini gerçekleştiren genellikle birleştirilmiş (pooling) ve tam olarak bağlı (fully-connected layers) katmanlar vardır.

Bir Evrişimsel Sinir Ağının bir örnek üzerinde aşağıdaki şekildeki gibi inceleyebiliriz (Şekil 3.7).



$$K = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

Şekil 3.7 Evrişimsel Sinir Ağı Örneği

Yukarıdaki örnekte sensör verisi 5x5x1 giriş sensör verilerimizi göstermektedir. Bir Evrişim Katmanının birinci kısmında evrişim işleminin gerçekleştirilmesinde yer alan elemente, noktalı olarak 3x3x1 matris olarak temsil edilen sensör verisi üzerine gelmiş filtre matrisidir (P). K, 3x3x1 matris olarak seçilmiştir. Çekirdek, adım uzunluğu nedeniyle 9 kez kayar, her seferinde K ile çekirdeğin üzerinde durduğu sensör verisinin P kısmı arasında bir matris çarpma işlemi gerçekleştirilir.

### **3.1.8. Uzun / Kısa Süreli Bellek (Long / Short Term Memory (LSTM))**

Uzun kısa süreli bellek (LSTM), verileri rassal zaman boyunca tanıyan bir tekrarlayan sinir ağı (RNN) modelidir [78]. Öğrenilen başarımlar saklandığında, kaydedilen veriler değiştirilmez. Tekrarlayan sinir ağları, nöronlar içerisinde ileriye ve geriye bağlanmalara müsaittir.

Uzun / Kısa Süreli Bellek algoritması, zaman serilerini sınıflandırmak, işlemek ve öngörmek için zamansal gecikmeler düşünüldüğünde oldukça uygundur. Zamansal uzunluğuna göreceli duyarsızlık, alternatif RNN'ler ve farklı uygulamadaki diğer öğrenme yöntemlerine kıyasla LSTM'ye bir avantaj sağlamaktadır.

LSTM'ler, uzun vadeli bağımlılık probleminden kaçınmak için açıkça tasarlanmıştır. LSTM'ler bilgiyi uzun süre hatırlarlar ve bu işlemi öğrenerek yaparlar. Tüm tekrarlayan sinir ağları, bir sinir ağının yinelenen modülleri zincirinin formuna sahiptir. Standart RNN'lerde, bu yinelenen modül, tek bir tanh katmanı gibi bir yapı ile yapılmaktadır.

Bir LSTM ağı, LSTM üniteleri içerir ve bu üniteler diğer ağ üniteleri yerine kullanılır. LSTM birimi, uzun / kısa zamansal verilerinin bilgilerini saklar. Bu işlemi yaparken, tekrarlanan verilerde herhangi bir etkinleştirmek işlevinin kullanılmamasıdır. Dolayısıyla, eski verilerden saklanan değerler tekrarlamalı şekilde değişmez ve zamansal olarak geri yayılımla eğitilir.

LSTM birimleri, genellikle birkaç birim içeren "bloklar" içinde uygulanmaktadır. Bu tasarım, derin sinir ağlarında tipiktir ve paralel donanımlarla yapılan uygulamaları kolaylaştırır.

## 3.2. Öznitelik Seçme Yöntemleri

### 3.2.1. Relieff Öznitelik Seçme Algoritması

Relieff, 1992'de Kira ve Rendell tarafından geliştirilen ve öznitelik etkileşimlerine duyarlı olan ve özellik seçimine filtre yöntemi yaklaşımı getiren bir algoritmadır [81]. Başlangıçta ayrık veya sayısal özniteliklerde ikili sınıflandırma problemlerine uygulama için tasarlanmıştır. Her bir öznitelik için bir puan hesaplar ve daha sonra öznitelik seçimi için en üst puanlama ile öznitelikleri seçer. Alternatif olarak, bu puanlar alt modellemeyi yönlendirmek için öznitelik ağırlıkları olarak da uygulanabilir. Öznitelik puanlaması, en yakın komşu çiftleri arasındaki öznitelik değeri farklılıklarının tanımlanmasına dayanır. Aynı sınıftaki komşu bir örnek çiftinde bir öznitelik değeri farkı gözlemlenirse, öznitelik puanı düşer. Alternatif olarak, bir komşu örnek çiftinde farklı sınıf değerlerine sahip bir öznitelik değeri farkı gözlenirse, öznitelik puanı artar.

Bilinen iki sınıfa ait,  $p$  özniteliklerinin  $n$  örneğini içeren bir veri kümesi aldığımızda; veri kümesinde, her öznitelik  $[0, 1]$  aralığında ölçeklendirilmelidir (ikili veriler 0 ve 1 olarak kalmalıdır). Algoritma  $m$  kez tekrarlanacaktır.  $P$  uzunluklu ağırlık vektörü ( $W$ ) değeri de sıfırdan başlatılır. Her yinelemede, rastgele bir örneğe ait olan öznitelik vektörünü ( $X$ ) ve her sınıftan  $X$ 'e en yakın olan örneğin (Öklid mesafesine göre) öznitelik vektörlerini alalım. En yakın aynı sınıf örneğine 'nearHit', en yakın farklı sınıf örneğine 'nearMiss' denildiğinde; ağırlık vektörünün eşitliği (3.10)'daki gibi olmaktadır:

$$W_i = W_{i-1} - (x_i - nearHit_i)^2 + (x_i - nearMiss_i)^2 \quad (3.10)$$

(3.10) eşitliğini incelediğimizde, herhangi bir özneliğin ağırlığı; aynı sınıfın yakındaki örneklerinde, diğer sınıfın yakındaki örneklerinden daha fazla olması durumunda, bu öznelikten farklı olması durumunda azalır ve ters durumda da artar.  $M$  yinelemelerden sonra, oluşan ağırlık vektörü alaka düzeyi vektörü olmaktadır.

### 3.2.2. Ki-Kare Öznitelik Seçme Algoritması

Öznitelik seçimi işlemi için; veri kümesinden en alakalı öznitelikleri çıkarma ve daha sonra modelin daha iyi performans göstermesi için makine öğrenme algoritmalarını uygulama işlemidir. Çok sayıda alakasız öznitelik, eğitim süresini üstel olarak arttırmakta ve fazla takılma riskini arttırmaktadır [82].

Ki-kare öznitelik seçme yöntemi istatistik tabanını kullanarak en alakalı özniteliği seçmek için yaygın olarak kullanılmaktadır [82]. Ki-kare öznitelik seçme algoritması ile her öznitelik değeri (3.11) eşitliği ile bulunmaktadır.

$$X^2 = \sum_{i=1}^m \sum_{j=1}^k \frac{(A_{ij} - \frac{R_i C_j}{N})^2}{\frac{R_i C_j}{N}} \quad (3.11)$$

Yukarıdaki formülde  $m$  verilen aralık sayısı,  $k$  sınıf sayısını vermektedir.  $R_i$  ve  $C_j$  sırasıyla,  $i$ nci aralıktaki örnek sayısını ve  $j$ nci sınıftaki örnek sayısını vermektedir.  $A_{ij}$   $j$ nci sınıfta ve  $i$ nci aralıktaki örnek sayısını,  $N$  ise veri kümesinde bulunan toplam örnek sayısını temsil etmektedir. (3.11) eşitliği ile her bir öznitelik için Ki-kare değeri hesaplanarak, öznitelik bazlı anlamlılık bulunur ve her bir öznitelik tek tek değerlendirilir.

### 3.2.3. Bilgi Kazancı Öznitelik Seçme Algoritması

Bilgi kazancı öznitelik seçme algoritması, veri kümesindeki her özniteliğe bakar, bilgi kazancını hesaplar ve sınıf etiketi ile ne kadar önemli ve alakalı olduğunu ölçer. Bir öznitelik için bilgi kazancının hesaplanması, tüm veri kümesi için sınıf etiketinin entropisinin hesaplanmasını ve bu özelliğin her olası değeri için koşullu entropilerin çıkarılmasını içerir. Entropi hesaplaması, sınıf etiketinin özellik değerine göre sıklıkta sayılmasını gerektirir. Daha ayrıntılı olarak, tüm örnekler öznitelik değeri ile seçilir, daha sonra bu sınıflardaki her bir sınıfın oluşum sayısı sayılır ve entropi hesaplanır. Bu adım, özelliğin olası her değeri için tekrarlanır. Bir alt kümenin entropisi, eğitim örneklerinin sınıf üyeliğini özellik değerine göre belirleyen bir sayım matrisi oluşturarak daha kolay hesaplanabilir [83].

Bilgi kazancı ile özniteliğin ne kadar alakalı olduğunu bulabilmemiz için önce Entropi'nin hesaplanması gerekmektedir. Daha sonra da bulunan Entropi'ye göre öznitelik bazlı bilgi kazançları hesaplanması gerekir. (3.12) eşitliği ile veri kümesinin Entropi'si hesaplanmaktadır:

$$Entropy = - \sum_{k=1}^K p_k \log_2 p_k \quad (3.12)$$

(3.12) eşitliğinde  $p_k$  değeri  $k$  ( $K = 1, \dots, k$ ) sınıfına ait örneklerin oranını gösterir. Bulunan Entropi'ye göre öznitelik bazlı bilgi kazançları da (3.13) eşitliği ile hesaplanmaktadır:

$$BilgiKazancı = Entropy - \frac{m_L}{m} Entropy_L - \frac{m_R}{m} Entropy_R \quad (3.13)$$

(3.13) eşitliğinde  $m$  toplam örnek sayısıdır,  $m_L$  ve  $m_R$  örnekleri  $k$  sınıfına aittir,  $K = 1, \dots, k$ .

#### 3.2.4. Kazanım Oranı Öznitelik Seçme Algoritması

Bilgi Kazancı yöntemi çok çeşitli değerlere sahip öznitelikleri seçme eğiliminde olduğundan sonuçları sapmalı çıkabilmektedir [11]. Kazanım oranı ile bu sapmanın azaltılması hedeflenmektedir. Sapmayı azaltmak için bölünme bilgisi (Split Information) kullanılmaktadır. Bölünme Bilgisi (3.14) numaralı eşitlikle gösterilmektedir.

$$BölünmeBilgisi_A(S) = - \sum_{i=1}^v \left( \frac{|S_i|}{|S|} \right) \log_2 \left( \frac{|S_i|}{|S|} \right) \quad (3.14)$$

Yukarıdaki (3.14) eşitliğindeki değer,  $S$  eğitim veri kümesini,  $S_i$ ,  $A$  özniteliğindeki bir testin  $v$  sonuçlarına karşılık gelen  $v$  bölümlerine bölerek oluşturulan bilgiyi temsil eder.

Kazanım Oranı, bilgi kazancı değerlerini, bölünme bilgisine oranlayarak bir çeşit normalizasyona tabi tutar [11]. Bölünme bilgisini kullanarak kazanç oranı (3.15) eşitliği ile tanımlanır:

$$KazancOranı(A) = \frac{BilgiKazancı(A)}{BölünmeBilgisi_A(S)} \quad (3.15)$$

Yukarıdaki (3.15) eşitliğindeki SplitInfo değeri (3.14) eşitliği ile hesaplanırken; Bilgi kazancı değeri (3.13) eşitliği ile hesaplanmaktadır.

## 4. ORTAM BAZLI FİZİKSEL AKTİVİTE TANIMA İÇİN VERİ TOPLAMA

### 4.1. Donanımların Kurulumu

FA sınıflandırmasında sensör yerleşimleri ile alakalı literatürde değişik çalışmalar mevcuttur. Bu çalışmaların amacı vücudun hangi yerine sensör takmanın FA sınıflandırmasına etkisinin daha yüksek olduğunu bulmaktır. İnsan vücuduna, takılabilir (inertial) bir ölçüm biriminin (IMU) yerleştirilmesi için ayağın, belin, uyluğun, ayak bileğinin, el bileğinin en sık tercih edilen yerler olduğu gösterilmektedir [55]. Buna karşılık, Mokhlespour vd. çalışmasındaki katılımcıların büyük bir kısmı (% 44,4) bir Short Sleeved (STS) kullanmak yerine vücutlarında herhangi bir giysiyi giymeyi tercih etmektedirler. Kısa kollu tişörtler, bileklikler, çoraplar, kolsuz tişörtler ve ayak bileği saatleri, bir STS için en sık seçilen giysilerdir. Mokhlespour vd. çalışmasına göre, tercih edilen yerler birbirine yakındır. Cep telefonlarının yaygınlığı, kullanım kolaylığı ve iletişim kanalları da bunları daha aktif bir yaşam tarzı için uygun bir platform haline getirmektedir. Bu çalışmada ölçüm birimi olarak cep telefonunun seçilmesinin nedeni, cep telefonunda birçok sensör bulunmasıdır. Ayrıca, sadece birçok sensörün dahil edilmesi ve kullanılabilirliği değil, aynı zamanda yaygın kullanım, sıfır kurulum maliyeti ve ergonomik kolaylık da cep telefonlarını kullanmamızı cazip hale getirmiştir. Vücut üstü sensörlerin kullanımıyla karşılaştırıldığında, cep telefonları kullanıcının hareketini kısıtlamaz, vücuda monte edilmesine gerek yoktur ve sensörlerin kurulumunu veya kalibrasyonunu gerektirmez. Vücudun çeşitli yerlerine sensörlerin takılmasıyla oluşturulan IMU sistemleri veya erişilmesi zor olan STS sistemleri insanların günlük yaşamda kullanması zordur. Bu nedenle, insanların, çoğu zaman yanlarında taşıdıkları cep telefonlarını kullanmanın FA tanıma için daha pratik bir çözüm olduğu düşünülmüştür.

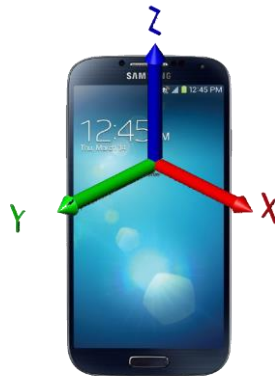
Aktivite verilerini toplarken Innovo oksimetre cihazı [56] ve iki Android tabanlı akıllı telefon [57, 58] kullanılmıştır. SpO<sub>2</sub> sensörü bir kişinin oksijen saturasyon seviyesini ölçmek için bir prob ile kullanılmıştır[56]. Sensör SpO<sub>2</sub> seviyesini saniyede beş kez okur. Aslında, oksijen doygunluğunu ölçmek için farklı cihazlar vardır. Bununla birlikte, parmak üzerinde kullanılan ölçüm cihazları daha yaygındır [59] ve insanlar bunları kolayca alabilir. Kullanılan akıllı telefonların seçilmesindeki amaç, içerisinde kararlı Android işletim sistemi barındırmaları, hesaplama için güçlü işlemcileri olması, yeterli bellek boyutu ve uzun pil ömrü gibi birçok avantajlarıdır. Ayrıca,

oksimetre sensörüne Bluetooth aracılığıyla ayrı olarak bağlanabilme özelliğine sahiptir. Şekil 4.1'de gösterilen oksimetre cihazı, laptop veya masaüstü bilgisayara Bluetooth ile bağlanır. SpO<sub>2</sub>'yi aşırı hassasiyetle ölçen ve saniyede 5 okuma yapabilen parmağa takılan bir prob'u vardır.



Şekil 4.1 Kullanılan Oksimetre Cihazı

Çalışmada kullanılan akıllı telefonlar ACC, MAG, GYR, CLI ve ortam sıcaklık sensörüne sahiptir. Çalışma sırasında, veri kümesini toplamak ve kaydetmek için bir mobil uygulama geliştirilmiştir. Her akıllı telefon 3 eksenli ACC, 3 eksenli GYR, 3 eksenli CLI ve 3 eksenli MAG sensörü içerir. Sensörlerden gelen veriler akıllı telefona etki eden hareketi Şekil 4.2'de gösterilen eksenler doğrultusunda ölçer. Her akıllı telefon 3 eksenli ACC, 3 eksenli GYR, 3 eksenli CLI ve 3 eksenli MAG içerir. Ham sensörlerden gelen veriler, akıllı telefon işletim sisteminden  $m/sn^2$  üç eksende elde edilir. Örneğin, ham ACC sensörü verileri  $Acc_i = \langle x_i, y_i, z_i \rangle, i = (1, 2, 3 \dots)$  vektörlerini içerir.



Şekil 4.2 Cep telefonu eksenleri

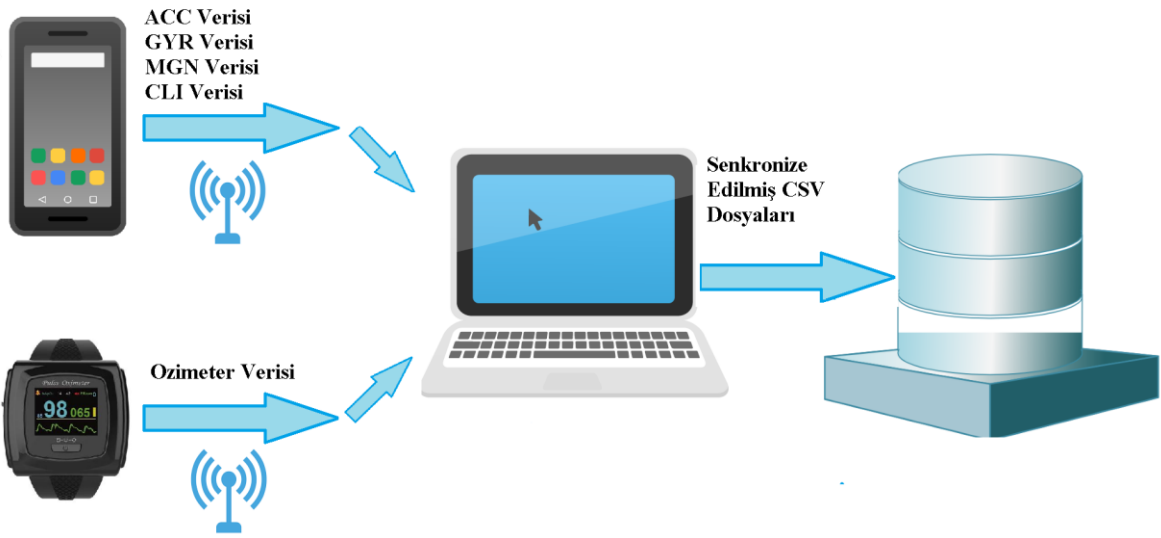
Sensörler farklı vücut pozisyonlarına yerleştirilir. Oksimetre sol el bileğine ve probu sol işaret parmağına takılmıştır. Bir sensör takmak için en iyi konumun ne olduğunu belirlemek için önemli miktarda literatür vardır. Cleland vd. [60], aktivitenin çeşitliliğini tespit etmek için veri sağlama yeri olarak kalçanın en iyi yer olduğunu bulmuşlardır. Ayrıca, sensörler arasındaki sınıflandırma doğruluğunda birbirine oldukça yakın olan farklar bulmuşlardır (sol kalça %97.81, göğüs %96.91). Esfahani ve Nussbaum [55], STS ve IMU'ların FA izleme için tercih edilen yerleştirme ve kullanılabilirliğini bulmayı amaçlamışlardır. Kısa kollu Tişört, bileklik, çorap, kolsuz Tişört ve ayak bileği kolunun bir STS için tercih edilen giysiler olduğunu göstermişlerdir. Boerema vd. [61] FA'yı ACC ile ölçmek için en uygun sensör yerleşimini ele almışlardır. Sensör konumunun, FA tipinden etkilendiğini bulmuşlardır. Ayrıca, farklı sensör pozisyonları uygulayan serbest çalışmaların karşılaştırılmasının imkânsız olduğunu belirtmişlerdir. Özdemir [62], düşme tespiti üzerine toplam 378 sensör yerleştirme kombinasyonunu araştırmıştır. Çalışmada sensör yerleşimine göre aktivite bulma doğruluğu verileri birbirine çok yakındır (bel %99.87, sağ uyluk %99.61).

Literatüre göre, sensör yerleşimine göre FA bulma doğruluğu verileri birbirine çok yakındır. Ayrıca, insanlar çantasında bir cep telefonu ile hareket etmediğinde, genellikle telefon el ile birlikte kullanılır veya kullanıcının pantolon cebinde bulunur. Bu nedenlerden dolayı, çalışmamızda bu iki yerleşim kullanılmıştır. Sensör yerleşimlerinin FA tanıma üzerindeki etkisini değerlendirmek için akıllı telefonlardan biri deneklerin pantolonunun sağ cebine konulmuştur ve sağ elinde ikinci bir akıllı telefonu tutturulmuştur.

Oksimetre ve akıllı telefon cihazlarından toplanan verileri Bluetooth üzerinden dizüstü bilgisayara aktarılmıştır. Sensörleri senkronize etmek için, kullanılan tüm sensörler, verilerin aktarıldığı bilgisayarla saat, dakika, saniye olarak aynı zamana ayarlanmıştır. Cep telefonu sensörlerinin örnekleme hızı 200Hz ve oksimetre cihazının örnekleme hızı 1Hz'dir. Bu yüzden oksimetre verileri her saniye için cep telefonu sensörü verisiyle eşleştirilmiştir. Bu şekilde, herhangi bir saniyedeki 200 cep telefonu sensörü verisine uyan oksimetrenin değeri aynıdır. Şekil 4.3, veri toplama ve transfer yapılandırmanın fiziksel görünümünü göstermektedir. Şeklin ilk bölümünde, veriler Bluetooth aracılığıyla cep telefonunun sensörleri ile toplanmıştır ve CSV dosyasına kaydedilmiştir. Benzer şekilde, şeklin ikinci



bölümünde, oksimetre sensöründen gelen verileri Bluetooth aracılığıyla toplanmıştır ve farklı bir CSV dosyasında saklanmıştır. Her denek için toplanan veriler farklı bir CSV dosyasında saklanmıştır. SpO<sub>2</sub> ile akıllı telefon sensörleri tek bir veri kümesi dosyasına göre hizalanmıştır ve depolanmıştır. Her dosya, 27 sütun içerir; bunlardan her biri deneklerden her 1 milisaniyede toplanan sensör değerlerini (öznitelikleri) gösterir (Çizelge 4.1). Fiziksel aktiviteleri tanımlamak için bizim tarafımızdan yapılan etiketler Çizelge 4.2'de gösterilmektedir (örneğin, açık alan yürüyüşü için sınıflandırma etiketi '0' dır). Her dosya satır başına 27 sütun içerir, sütun içeriği Çizelge 4.3'de gösterilmiştir.



Şekil 4.3:Veri toplama mimarisi

Çizelge 4.1 Sensörlerin vücutta yerleşme yerleri ve adetleri (örn. Elde tutulan telefonda 3 adet ACC sensörü vardır)

Sensor	Sensör Yerleşimi		
	El	Pantolon Cebi	Parmak
ACC	3	3	-
GYR	3	3	-
MAG	3	3	-
CLI	3	3	-
Hava Derecesi	-	1	-
SpO <sub>2</sub>	-	-	1
Nabız	-	-	1

Çizelge 4.2 Tanımlanan Fiziksel Aktivite Sınıfları

Aktivite	Süre (dakika)	Aktivite	Süre (dakika)
İç Alan Yürüme	10	Dış Alan Yürüme	10
İç Alan Durma	10	Dış Alan Durma	10
İç Alan Oturma	10	Dış Alan Oturma	10
İç Alan Koşma	10	Dış Alan Koşma	10

Çizelge 4.3 Toplanan sensör verilerinin bazı istatistiksel değerleri

	Minimum	Maksimum	Ortalama	Standart Sapma
Accelerometer Hand X	-23,26	28,19	0,03	2,04
Accelerometer Hand Y	-29,26	25,87	-0,06	2,12
Accelerometer Hand Z	-25,12	25,86	0,16	2,85
Gyroscope Hand X	-8,73	7,86	0,02	0,67
Gyroscope Hand Y	-8,72	7,53	0,05	0,62
Gyroscope Hand Z	-8,73	7,86	0,02	0,67
Magnetometer Hand X	-77,10	101,10	-1,29	20,63
Magnetometer Hand Y	-69,36	46,74	-21,29	16,83
Magnetometer Hand Z	-75,78	60,42	-11,58	14,85
Clinometer Hand X	-180,00	180,00	7,96	127,40
Clinometer Hand Y	-89,63	89,07	-24,10	14,48
Clinometer Hand Z	-179,97	179,96	-9,09	22,64
Accelerometer Trousers X	-26,39	28,51	0,05	2,86
Accelerometer Trousers Y	-26,10	28,62	-0,11	3,28
Accelerometer Trousers Z	-27,60	25,85	-0,06	2,67
Gyroscope Trousers X	-10,01	9,02	-0,01	0,88
Gyroscope Trousers Y	-10,02	10,00	-0,06	1,13
Gyroscope Trousers Z	-10,01	9,02	-0,01	0,88
Magnetometer Trousers X	-45,18	54,00	12,41	18,44
Magnetometer Trousers Y	-46,56	58,26	22,62	19,29
Magnetometer Trousers Z	-52,08	68,40	-10,27	18,90
Clinometer Trousers X	-180,00	180,00	25,41	92,20
Clinometer Trousers Y	-89,17	89,93	40,61	29,70
Clinometer Trousers Z	-180,00	179,99	53,64	60,45
Isı	11,15	40,22	25,71	4,08
Oksimetre	81,00	99,00	95,98	1,63
Nabız	44,00	240,00	84,81	19,78

## 4.2. Akıllı Telefon Sensörleri

Dahili akıllı telefon sensörleri, cihazların çevredeki bilgileri algılamasını sağlar. Bu sensörlerin kullanımı, kullanıcının fiziksel aktivitesini bulmak için kullanılacak ham verilerin toplanmasına izin verir. Bir akıllı telefon olarak tipik bir cihaz, çoklu ölçümlerle etkileşime girebilir, bu nedenle farklı sensörlerden oluşan bir koleksiyon kullanmayı gerektirir. Hem iOS hem de Android olmak üzere her akıllı telefon, ACC, GYR, MAG, parlaklık, kamera, mikrofon vb. gibi birçok farklı sensör ile donatılmıştır. Bu çalışmada Android akıllı telefonunda bulunan en yaygın dört farklı sensör dikkate alınmıştır. Bunlar, ACC, GYR, MAG ve CLI sensörüdür. Aşağıda, kullandığımız, değerlendirilen her sensörün yapısının kısa bir tanımı açıklanmıştır. Android şu ana kadar en yaygın kullanılan akıllı telefon işletim sistemi olduğu için [63], bu araştırmada yalnızca Android akıllı telefonlar kullanılmıştır.

ACC, cihazın hızlanmasını üç eksen olan x, y ve z boyunca ölçmeye izin veren ve  $m/s^2$  cinsinden ölçümler alabilen bir sensördür. ACC sensörünün ana uygulaması, hareket algılamayı, aygıt uyumunu ve hareket değerlendirmesini içerir. Tahmin edilebileceği gibi, bu sensör hareket bağlamını ilgilendiren her durumda yer almaktadır [63].

GYR, cihaz oranının üç eksen x, y ve z boyunca dönüşünü ölçmeye izin verir ve rad/s cinsinden ölçümler alabilir. Bu cihaz, cihaza etkilenen kesin hareket miktarını daha iyi değerlendirmek için ACC sensörü ile birlikte kullanılabilir. Bu sensör, titreşim kütleleri, ivme ve dolayısıyla yön değişiminin metal yüzeyleri birleştirerek gerçekleştirilir. Bu hareketin neden olduğu titreşim miktarını ölçerek tespit edilebilir [63].

MGN, Tesla'daki x, y ve z eksenleri arasındaki yoğunluğunu ölçen çevresel manyetik alanı algılar. Bu sensörün içinde bulunduğu ana uygulama, cihazın Dünya jeomanyetik alanı ile ilgili oryantasyonunun edinilmesi ile ilgilidir. Bu prosedür, GPS modülü [63] ile birlikte kullanıldığında cihazın harita üzerinde doğru konumlandırılmasını sağlayabilir.

CLI, iki orijinal tek tip Fiber Bragg Izgarası (FBG), küresel bir kütle, üst kısmı kesik olan bir silindirik kırıntıdan oluşur. FBG'ler, iletişimin yanı sıra optik sensörler alanında da yoğun araştırmalar yaptılar. Sıcaklık, zorlanma, basınç, kuvvet, burulma, eğim açısı vb. gibi parametreleri ölçmek için çeşitli FBG sensörleri geliştirilmiştir [63].

### 4.3. Denekler

Veri toplamaya toplam 12 denek, 6 erkek ve 6 kadın katılmıştır. Denekler, çalışanlar, öğrenciler veya emeklilerdir. Veri toplama işlemi 2017 sonbaharında Ankara ilinde gerçekleştirilmiştir. Bu çalışmaya alınan deneklerin özellikleri Çizelge 4.4'de gösterilmiştir.

Çizelge 4.4 Deneklerin özellikleri

Değişkenler	Ortalama	Min-Maks
Yaş (yıl)	35	20-65
Boy (cm)	175	155-185
Kilo (kg)	75	55-130
Vücut Kitle Endeksi (kg/m <sup>2</sup> )	22	18-30
Nabız	80	50-200
SpO <sub>2</sub> (%)	95	83-99

### 4.4. Veri Toplama Protokolü

Her biri 10 dakika süren dört temel etkinlikten oluşan bir protokol tanımlanmıştır. Bu aktiviteler yürüyüş, koşma, oturma ve ayakta durmadır. Pengfei vd. [64], çalışmalarında bunların yaygın olan faaliyetler olduğunu söylemişlerdir. Aynı faaliyetler iç ve dış mekânlarda da gerçekleştirilmiştir. Her bir deneğin bu protokolü izlemesi, ayrıca tanımlanmış zamanda ve yerde açıklanan tüm FA'ları gerçekleştirilmesi gerekmektedir. Denekler FA'ya başlamadan önce, onlar için talimatlar hazırlanmıştır. Bu talimatlar, deneklerin bu aktiviteleri nasıl gerçekleştireceğini açıklamaktadır. Örneğin, bu talimatlar arasında, bir denek kaç dakika yürüyecek, koşacak, oturacak ya da ayakta duracak bilgileri bulunmaktadır. Ayrıca, bu etkinlikleri nerede yapmaları gerektiği bilgisi de bulunmaktadır (açık ya da kapalı alanda). Her deneğin bu talimatları izlemesi gerekmektedir. Ayrıca FA sırasında deneklerle bir araya gelinerek, herhangi bir aksaklık, öngörülemez olaylar veya hatalı durum hemen ele alınmıştır.

Sonuç olarak, kaydedilen veri kümesi 4 farklı FA'dan ve 2 farklı konumdan veri içermektedir. Protokoldeki bütün FA'lar yaklaşık 10 dakika süreyle gerçekleştirilmiştir. Veri toplama sırasında, 2 farklı konum için 4 farklı FA'dan biri olarak etiketlenen 16 saatten fazla veri toplanmıştır. Tüm on iki erkek ve kadın denekler, tüm etkinlikleri gerçekleştirmiştir. Ayrıca, gün boyunca farklı zamanlarda veri toplama gerçekleştirilmiştir. Bu saatler sabah 07:00-11:00, öğleden sonra

12:00-16:00 ve akşamları 18:00-00:00 arasındadır. Tüm denekler, etkinliği gerçekleştirmeden önce etkinliğin nasıl gerçekleştirileceği hakkında bilgilendirilmiştir. Ayrıca, deneklerin gizlilik sorunları bu kişilerden etik onay alınarak halledilmiştir. 27 özniteliğimizin bir sınıf değeri vardır. Her kayıt ACC, GYR, MAG, CLI ve oksimetre sinyallerinden oluşuyordu ve FA etiketi ile etiketlenmiştir. Toplanan tüm kayıtlar, Grup A (Grup 18 - 34 yaş arası), Grup B (35 - 54 yaş arası) ve Grup-C (yaş 55 - 65 yaş arası) olmak üzere 3 farklı yaş grubuna ayrılmıştır. Her denek bütün FA'ları gerçekleştirilmiştir ve kayıtlar 1 milisaniyelik aralıklı olarak etiketlenmiştir. Bu amaçla, android cihazda kullanılmak üzere bir mobil uygulama geliştirilmiştir. Bu uygulama, etkinlik gerçekleştirildiğinde sensör verilerini CSV dosyası olarak kaydeder. Ardından verileri Bluetooth ile dizüstü bilgisayara aktarılmıştır. Bu arada, CSV dosyasının adı etiket ve denek ile değiştirilmiştir. Örneğin, a1\_outdoor\_walking, a2\_indoor\_standing, vb. Her denek için toplanan veriler farklı bir CSV dosyasında saklanır. Her denek için bu verilerden oluşan 8 farklı dosya vardır. Bu dosyaların adları ayrıca etiketleri de taşır.

Farklı yaş gruplarını kullanmanın amacı, yürüme ve koşu hızı farklılıklarını yaşa göre ortadan kaldırmaktır. Yürüme hızı aralığı olarak 0,75 ile 1,75 m/s kullanılmıştır. Otter vd. [65] bu hızı normal yürüme hızı olarak belirlemişlerdir. Ayrıca koşu hızı olarak 1,75'ten büyük değerler seçilmiştir. Yürüme ve koşma hızını hesaplamak için, adım boyunca deneklerin ayakları arasındaki mesafe metre skalasıyla ölçülmüştür. Ardından, adım sayısı sayılmıştır. Daha sonra da, denekler 1 dakika yürümüştür ve attıkları adım sayısı sayılmıştır. Bu şekilde, deneklerin 1 dakika içinde kat ettiği toplam mesafe bulunmuştur. Son olarak, bu değer 60'a bölünerek saniyede kat edilen mesafe bulunmuştur. FA sırasında deneklerle bir araya gelinmiştir. Yürüme hızı aralığında yürümemişlerse, bu FA'yı tekrar etmeleri sağlanmıştır. Deneklerimizin yürüme ve koşu hızları Çizelge 4.5'de gösterilmektedir.

FA tanımayı değerlendirmek için geçmişte birkaç veri kümesi kullanılmıştır. Bu tez çalışması kapsamında oluşturulan veri kümesi ile mevcut diğer veri kümelerinin karşılaştırması Çizelge 4.6'da sunulmaktadır. Bu tablo, bu veri kümelerinin özelliklerinden bazılarını özetlemektedir. Çok sayıda örnek ve MAG, CLI ve oksimetre sensörlerinin kullanımı veri kümemizin en önemli noktalarıdır.

Çizelge 4.5 Deneklerin yürüme ve koşma hızları

Denekler	Yürüme Hızı (m/s)	Koşma Hızı (m/s)
<b>A1</b>	1,39	1,86
<b>A2</b>	1,42	1,82
<b>A3</b>	1,35	1,92
<b>A4</b>	1,40	1,99
<b>B1</b>	1,54	2,06
<b>B2</b>	1,32	1,93
<b>B3</b>	1,47	2,05
<b>B4</b>	1,30	1,91
<b>C1</b>	1,29	1,82
<b>C2</b>	1,22	1,84
<b>C3</b>	1,40	1,95
<b>C4</b>	1,19	1,83

Çizelge 4.6 Karşılaştırma için kullanılan veri kümelerinin karakteristikleri

Veri Kümesi	Aktivite Sayısı	Örneklemedeki Kişi Sayısı	Sensör Yeri	ACC	GYR	MAG	CLI	Oximeter	Örnekleme Oranı	Örnekleme Sayısı
WISDM v1.1 [66]	6	29	Ön Pantolon Cebi	✓	X	X	X	X	20Hz	1098207
Daphnet FoG [67]	2	10	Gövde, uyluk, ayak bileği	✓	X	X	X	X	64Hz	1917887
Skoda [68]	10	1	Kollar (20 pozisyon)	✓	X	X	X	X	98Hz	701440
UCI HAR [25]	6	30	Bel	✓	✓	X	X	X	50Hz	10299
Kendi Veri kümemiz	8	12	Sağ pantolon cebi, sağ el, sol işaret parmağı	✓	✓	✓	✓	✓	200Hz	13959902

## 5. SENSÖR ANALİZLERİ

Hangi sensör verilerinin daha anlamlı olduğunu bulmak için sensör bazlı analizler yapılmıştır. Bu analizler tek kipli ve çok kipli olarak yapılmıştır. Tek kipli analizlerde sadece tek sensör verisinden gelen veriler kullanılırken; çok kipli analizlerde birden fazla sensör verisinin füzyon edilmiş hali kullanılmıştır. Her görev için (fiziksel aktivite tanıma, uyku apnesi tanıma, vb.) farklı çok kipli senaryolar oluşturulmuştur. Bu görevlerdeki tek kipli kullanımlarda ise sadece tek sensör verisinden gelen veriler kullanılmıştır.

Bu amaçla fiziksel aktivite tanıma işlemi için üç farklı çok kipli analiz ve sekiz farklı tek kipli analiz tanımlanmıştır. Çok kipli analizlerde, birleştirme yöntemi olarak 5.2. bölümünde anlatılacak olan füzyon operatörü kullanılmıştır. İlk olarak, hem el hem de pantolon cebinden toplanan cep telefonu sensörlerinin öznelikleri füzyon edilmiştir. Daha sonra, oksimetre, termometre ve kalp atış hızı sensörleri da akıllı telefon sensörleri ile füzyon edilmiştir. İkinci çok kipli özneliklerimiz için, pantolon cebindeki cep telefonu verileri ile oksimetre, termometre ve nabız verileri füzyon edilmiştir. Son olarak da, elde taşınan cep telefonu verileri ile oksimetre, termometre ve nabız verileri füzyon edilmiştir.

Değişik fiziksel aktiviteleri sınıflandırmak için birbirinden ayrı olan sensörler kullanılmıştır. Bu sensörler elektrokardiyografi (EKG), ACC, MAG ile GYR şeklindedir. Kullanılan bu dört farklı sinyal öznelik seviyesinde füzyon edilmiş ve bu sayede çok kipli model kullanılmıştır. Sınıflandırma methodu olarak Destek Vektör Makineleri, k En Yakın Komşu, Naive Bayes ile Rastgele Ağaç algoritmaları kullanılmıştır. Füzyon yöntemi ile alakalı önemli sonuçlar sunulmuştur. Bu çalışmaları test etmek için MHealth veri kümesi kullanılarak deneysel çalışmalar yapılmıştır.

Ayrıca çok kipli yaklaşımın, iyi bilinen öğrenme yöntemleriyle birlikte uyku apnesi probleminin tespit doğruluğu üzerindeki etkisi araştırılmıştır. Spesifik olarak, OSA sınıflandırması için EKG ile SpO<sub>2</sub> gibi iki fizyolojik sinyalin öznelik düzeyinde füzyonunu gerçekleştirilmiştir; NB (Naive Bayes), SVM ve kNN sınıflandırıcıları için kullanılan füzyon stratejisine ilişkin kapsamlı bir yaklaşım kullanılmıştır. Uyku apnesi çalışması için oksijen doygunluğu ve elektrokardiyogram sensörleri ile Relief

öznitelik azaltma methodu yapılmıştır. Daha sonra NB, kNN ve SVM sınıflandırıcıları öğrenme algoritması olarak tasarlanmıştır.

Yapılan analizler sayesinde tek kipli ya da çok kipli olarak sınıflandırmanın başarımı test edilmiştir. Ayrıca kullanılan öznitelik seçme ve füzyon yöntemleri ile de sınıflandırmanın başarımının ne kadar arttığı araştırılmıştır.

### 5.1. Öznitelik Seçimi Yöntemi

Elimizdeki sensör verileri kullanarak sınıflandırma algoritmasının sonuçlanmasına katkı yapacak yöntemler geliştirilmiştir. Bu amaçla yeni özniteliklerin bulunması ve veri kümesinin küçültülmesi amacı ile öznitelik seçimi yöntemleri kullanılmıştır. Bu yöntemlerin kullanılması esnasında filtreleme tabanlı ki-kare, bilgi kazancı, kazanç oranı ile ReliefF methodları seçilmiştir. Daha sonra kullanılan öznitelik seçimi yöntemi ile oluşturulan özniteliklere göre puan sıralamaları belirlenmiştir. Elde ettiğimiz puan sırasına göre en uygun sayıda öznitelik içeren veri kümeleri ayrıca tespit edilmiştir. Kullanılan dört methoda göre seçilen özniteliklere baktığımızda oluşturulan veri kümeleri seçilen özniteliklere göre bakıldığında, seçilmiş özelliklerin hepsinin benzer öznitelikler ile oluştuğu görülmüştür.

Öznitelik seçim işlemi, veri kümemiz ile oluşturulan yeni füzyon edilmiş verilerimiz için tekrarlanmıştır. Buradan bulunan sıralamalara dikkate alınarak çok kipli olarak füzyon methodu yapılmıştır. Örnek olarak baktığımızda, bilgi kazancı yöntemi ile öznitelik seçimi yapıldığında ortaya çıkan değerlere göre en yüksekler sırasıyla füzyon edilmiştir. Bu yöntemler ile bulunan seçilip füzyon edilmiş ve seçilmeden kullanılan sensör verilerine ait veri kümeleri Çizelge 5.1’de gösterilmektedir.

Çizelge 5.1 Öznitelik seçimine göre füzyon edilmiş ve tek kipli olarak oluşturulan senaryolar

Senaryo Numarası	Model Tipi	Kullanılan Öznitelik	Boyutu
Senaryo 1	Çok Kipli	İvme Ölçer, jiroskop, manyetometre, EKG	23
Senaryo 2	Çok Kipli	İvme Ölçer, jiroskop,	21
Senaryo 3	Çok Kipli	İvme Ölçer, jiroskop	15
Senaryo 4	Tek Kipli	İvme Ölçer	9
Senaryo 5	Tek Kipli	Jiroskop	6
Senaryo 6	Tek Kipli	Manyetometre	6
Senaryo 7	Tek Kipli	EKG	2



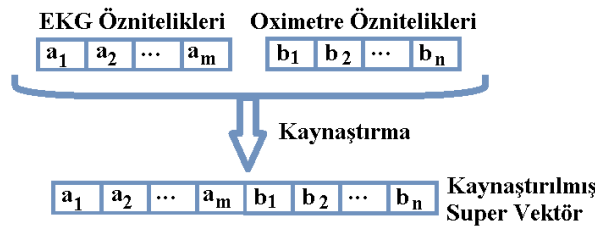
## 5.2. Öznitelik Füzyonu

Sensör sinyallerine dayanan öznitelik seviyesi füzyonunun amacı, sınıflandırıcının eğitiminden önce bu sinyal niteliklerini birleştirerek her sensör sinyalinin yakaladığı ayrıntıları kullanmaktır. Sınıflandırma algoritması, eğitiminden önce öznitelikleri füzyon ederek her verinin içerisinde barındırdığı çeşitli bilgilerden faydalanılmıştır. Toplam ve maksimum havuzlama gibi farklı kaynakların verilerini füzyon için değişik methodlar bulunmaktadır. Bu yöntemlerde dezavantajlar bulunabilir. Örneğin toplam havuzlama yöntemleri, sensör verileri içerisindeki gerekli bilgiyi kaybedebilirler. Bu problemten dolayı, sensör verileri ile kullanılan anlamlı bilgileri bulabilmek için füzyon yöntemi kullanılmıştır.

Öznitelik düzeyinde füzyon esnasında, sensör verilerimizin oluşturduğu vektörleri toplu olarak karşılayabilmek için bir tane vektör oluşturulmuştur.  $f_i = [a_1, a_2 \dots a_n]$  ve  $f_j = [a_1, a_2 \dots a_m]$  sensör verilerimizin öznitelik vektörleri olsun ve n ve m bu vektörlerin ilgili boyutunu temsil etsin. Bu öznitelik vektörlerinde  $a_k$  notasyonu bir özellik vektörünün k'ncı özneliğini temsil eder. Bu öznitelik vektörleri kullanılarak (5.1) deki öznitelik süper vektörü oluşturulmaktadır. Bu süper vektörün boyutu n+m boyutundadır.

$$f_s = [a_{i_1}, a_{i_2}, \dots, a_{i_n}, a_{j_1}, a_{j_2}, \dots, a_{j_m}] \quad (5.1)$$

Bu süper vektör yöntemini kullanarak, EKG ile SpO<sub>2</sub> verilerini kullanma esnasında, bu verilere ait vektörler füzyon edilmiştir. Sensör verisi düzeyinde füzyon esnasında, EKG ve SpO<sub>2</sub> sensörlerinin vektörleri son şekli verilen özniteliklerin temsili amacıyla tek bir vektör olarak füzyon edilmiştir. Bu veri füzyon methodunun şeması Şekil 5.1'de gösterilmiştir. Bu şemada,  $a_i$  EKG öznitelik vektörünü göstermektedir.  $b_j$  ise SpO<sub>2</sub> öznitelik vektörünü göstermektedir. Bu vektörler de m ve n uzunluğundadır. Bu method basitliği ve füzyon işlemi sonrasında daha anlamlı sonuçlar oluşturması sebebiyle geçmiş çalışmalarda da sıkça kullanılmaktadır [84].



Şekil 5.1 Öznitelik füzyonu

### 5.3. Sınıflandırıcı Tasarımı

Sınıflandırma işlemleri için farklı sınıflandırma algoritmaları kullanılmıştır. Bu algoritmalar, Kullanılan Temel Yöntemler Bölümü'nde anlatılan NB, kNN, SVM, Rastgele Ağaç ve Bagging RepTree şeklindedir. Bu sınıflandırma yöntemleri için farklı modeller oluşturulmuş, birbirleri içerisinde sınıflandırma başarımının karşılaştırması oluşturulmuştur. Öznitelik seçme, sensör pozisyonu belirleme gibi işlemlerde kullanılan bilinen algoritmalar daha hızlı sonuç verdiği için tercih edilmiştir. Burada bulunan sonuçlar ile gelecek çalışmalara taban oluşturulmuştur. Sınıflandırma işlemleri için Weka Yazılımı, 3.6.13 versiyonu kullanılmıştır. Weka, grafiksel kullanıcı arayüzleriyle birlikte veri analizi ve tahmine dayalı modelleme için görselleştirme araçları ve algoritmalar koleksiyonu içermektedir. Java programlama dili ile yazıldığı için platform bağımsızdır ve bütün işletim sistemleri ile kolaylıkla çalışabilir durumdadır. Bu faydaları ile çok büyük veri kümelerinde, kişisel bilgisayarlar ile yapılamayan işlemlerin, Linux sunucusu üzerinde yapılmasına olanak sağlamıştır. Ayrıca bu tez çalışması kapsamında, Weka ile yapılamayan, ayrı bir Java yazılımı ile sınıflandırma işlemi ihtiyacı durumlarında, Weka'nın Java için hazırlanmış olduğu weka.3.6.13.jar dosyası alınarak, geliştirilen Java uygulamasına entegre edilmiştir.

SVM methodu kullanıldığında farklı çekirdek fonksiyonları kullanılmıştır. Bu fonksiyonlardan radyal taban (RBF) ve polinomial çekirdek tek tek kullanılmıştır. Naive Bayes algoritması kullanıldığında Gaussian Dağılımı seçilmiştir. kNN algoritması kullanıldığında  $k=3$  kullanılmıştır. kNN algoritmasında uzaklıkları bulmak için Öklid yöntemi seçilmiştir. Topluluk methodları olarak Rasgele Ağaç kullanıldığında ağaç sayısı olarak 100 seçilmiştir. Bu seçim deneysel çalışmalar yapılarak bulunmuştur. Bagging RepTree kullanıldığında çanta sayısı 100 seçilmiş olup iterasyon 10 olarak belirlenmiştir. Testler yapılırken 10 kat çapraz doğrulama methodu ile yapılmıştır.

Bir diğer önemli eğitim kavramı, modelin bir sınıflandırıcı olarak ne kadar iyi performans gösterebileceğini değerlendirmek için kullanılan çapraz doğrulama fikridir. Bu amaçla, modeli eğitmek için veri alt kümesini seçtik ve sonra kalan veriler üzerinde, yani eğitimde kullanılmayan verileri test edilmiştir. K-katlama çapraz doğrulamada, eğitim kümesi  $k$  alt grubuna ayırır ve her alt kümeyi test kümesi olarak kullanırız. Diğer  $k - 1$  alt kümeleri birleştirilir ve eğitim kümesi olarak kullanılır. Eğitim

k kez tekrarlanır. Bu şekilde, eğitimden elde edilen modelin test aşamasında ne kadar iyi davranabileceğini değerlendirebiliriz.

#### **5.4. Kullanılan Veri Kümeleri**

##### **5.4.1. HPAwIO Veri Kümesi**

Bölüm 4'te ayrıntılı olarak anlatılan tez çalışması kapsamında toplanan veri kümesi kullanılmıştır. Veri kümesi 6 erkek ve 6 kadın içermektedir. Akıllı telefonlar içerisinde bulunan ACC, MAG, GYR, CLI ve ortam sıcaklık sensörleri ile oksimetre sensörü içeren veri kümesidir.

##### **5.4.2. MHealth Veri Kümesi**

Bu kullanılan veri kümesinde, ACC, GYR, MAG, EKG sinyallerinin yer aldığı MHealth veri kümesi [87] seçilmiştir. İlgili veri kümesi, fiziksel aktivite kayıtları ve bazı yaşamsal sensör verileri içeren 10 kişi barındıran kayıtlar vardır. Sensörler, sırasıyla kişinin göğüs, sağ bilek ve sol ayak bileklerine yerleştirilmiştir. Birden fazla sensör kullanılması, çeşitli vücut lokasyonlarının yaşadığı hareketleri, yani hızlanma, dönüş hızı ve manyetik alan yönelimi ölçmektedir. Böylece vücut dinamiklerini daha iyi yakalamaktadır. Göğüste konumlandırılan sensör, 2-derivasyonlu EKG ölçümleri de sağlar. Bu bilgi, temel kalp izleme, çeşitli aritmileri kontrol etme veya egzersizin EKG üzerindeki etkilerine bakmak için kullanılabilir. Algılama modaliteleri, 50 Hz'lik bir örnekleme hızında kaydedilmiştir. Aktiviteler, laboratuvar ortamında, konunun yürütülmesinde herhangi bir kısıtlama olmaksızın toplanmıştır, ancak denekler en iyi sonucu almayı denemişlerdir.

##### **5.4.3. Physionet Sleep Apne Veri Kümesi**

EKG ve SpO<sub>2</sub> sinyallerini içeren Physionet Sleep Apne Veritabanı [95] seçilmiştir. İlgili veritabanı, 35 kayıttan oluşan bir eğitim kümesine (a 01 ila a 20, b 01 ila b 05 ve c 01 ila c 10) bölünmüş 70 kayıttan ve 35 adet kayıt içereden bir test verikümesinden (x 01 ila x 35) oluşmaktadır. Her kayıt 7 ila 10 saat uzunluğundadır ve sürekli bir sayısallaştırılmış EKG sinyali, bir dizi apne ek açıklaması (aynı anda kaydedilen solunum ve ilgili sinyaller temelinde insan uzmanları tarafından türetilir) ve bir makine tarafından üretilen QRS ek açıklamaları içerir ( türünden bağımsız olarak tüm atımların normal olarak etiketlendiği). Ek olarak, sekiz kayda (a01 ila a04, b01 ve c01 ila c03) eşlik eden dört ek sinyal (Resp C, Resp A, göğüs ve abdominal

solunum eforu sinyalleri ile endüktans pletismografisi; Resp N, burun termistörleri kullanılarak ölçülen oronasal hava, SpO<sub>2</sub> ve oksijen satürasyonu) eklenmiştir.

Veri kümesi, her bir kaydın hem EKG hem de SpO<sub>2</sub> sinyallerinden oluştuğu ve uzmanlar tarafından apne veya oksijen doygunluğuna ve solunumuna bağlı normal olaylara açıklandığı 35 uyku çalışmasını içerir. Tüm kayıt üç farklı gruba ayrılmıştır. Grup-A yüksek uyku apnesi olan yirmi bütün gece uyku verisinin kayıtlarını, Grup-B düşük uyku apnesi olan beş bütün gece uyku verilerinin kayıtlarını ve Grup-C apnesi olmayan on bütün normal uyku gecesini içermektedir. Tüm kayıtlar, her dakika için uzmanlar tarafından işaretlenen, apnenin varlığını veya yokluğunu gösteren ek açıklamalar içerir.

Veri kümesi, tanı belirtisi için uyku apne hipopne indeksi (AHI) kullanır. Bu indeks uyku apnesi şiddeti için genel bir tanı sunar. AHI, saatteki uyku apnesi olaylarının sayısı olarak hesaplanır. Uyku apnesi tanısı tipik olarak 5-15 / saat = hafif; 15–30 / sa = orta; ve > 30 / h = şiddetli şeklindedir. Physionet veri kümesinin puanlama yöntemi, kayıtların her dakikası için “Apne dakika” veya “Normal dakika” şeklinde ek açıklamalar sağlar.

#### **5.4.4. Statlog Veri Kümesi**

Kalp krizi modeli çalışması için Statlog (Hearth) veri kümesi kullanılmıştır. Bu veri kümesi kalp hastalığı tahmin etmede kullanılabilen 13 tane değişkeni barındıran bir veri kümesidir. 270 tane örneklem barındırmaktadır. Bu veri kümesi 13 öznitelik içermektedir. Öznitelik bilgileri farklı insan sinyalleridir (yaş, cinsiyet, göğüs ağrısı tipi, istirahat kan basıncı, serum kolesterol (mg/dl), açlık kan şekeri >120 mg/dl, istirahat elektrokardiyografik sonuçları, maksimum kalp atış hızı, egzersize bağlı anjin, istirahate göre egzersizle tetiklenen ST değeri, pik egzersizi ile tetiklenen ST segmentinin eğimi, flourosopiyle renkli ana damarların sayısı).

#### **5.4.5. MIT-BIH Veri Kümesi**

Kalp krizi modeli çalışması için MIT-BIH Aritmi Veritabanı veri kümesi kullanılmıştır. MIT-BIH Aritmi Veritabanı, 1975 ve 1979 yılları arasında BIH Aritmi Laboratuvarı tarafından incelenen 47 kişiden elde edilen, iki kanallı EKG kayıtlarının 48 saat alıntısı içermektedir. 24 saatlik bir setten 23 adet kayıt seçilmiştir. Boston'daki Beth İsrail Hastanesinde yatan hastalar (% 60'ı) ile ayaktan hastalar (% 40'ı) karışık bir şekilde toplanmıştır. Bütün veri kümesindeki toplanan kayıtlardan, yatarak ve

ayaktan tedavi edilenler için EKG kayıtları anlamlı aritmileri içerecek şekilde seçilmiştir.

## **5.5. Deneysel Çalışmalar**

### **5.5.1. HPAwIO Veri Kümesi Çalışmaları**

#### **5.5.1.1. Kan Oksijen Doygunluğu ile Açık ve Kapalı Alan Sınıflandırılması Bulguları**

SpO<sub>2</sub> sinyalinin açık ve kapalı alanlardan nasıl etkilendiğini ve bu fizyolojik sinyalin kişinin kapalı alandaki yerini tespit etmek için kullanılıp kullanılmayacağını araştırmak hedeflenmiştir. Bildiğimiz kadarıyla, SpO<sub>2</sub> sinyali daha önce literatürde kişinin açık ve kapalı alan konumunun tespiti için kullanılmamıştır. Bu amaçla, SpO<sub>2</sub> sinyalini kullanarak bir veri toplama işlemi gerçekleştirilmiştir ve açık ve kapalı alan sınıflandırması için topluluk içi öğrenme tabanlı Rastgele Orman (RF) yöntemi önerilmiştir. Yöntemin etkinliğini değerlendirmek için, önerilen veri kümemizdeki seçili makine öğrenme algoritmaları olan Naive Bayes (NB), kNN ve Destek Vektör Makinesi (SVM) ile karşılaştırmalar yapılmıştır.

Bu çalışma kapsamında veri kümesi olarak 4. bölümde ayrıntılı olarak anlatılan veri kümesi kullanılmıştır. Kullanılan sınıflandırıcının etkinliği, büyük miktarda veri ile çalışırken gerçek zamanlı uygulamalar için önemlidir. Ağaç bazlı sınıflandırıcılar, hesaplama etkinlikleri nedeniyle bu gibi durumlarda tercih edilebilir. Önceki çalışmalar, torbalama ve güçlendirme gibi topluluk yöntemlerinin ağaç sınıflandırıcıları için etkili eğitim yöntemleri olduğunu göstermiştir. Özellikle, bir topluluk yöntemi olarak, RF'nin diğer topluluk yöntemlerinden daha üstün performans sağladığı gösterilmiştir. Bu nedenle, RF yöntemini sınıflandırıcı olarak kullanmak seçilmiştir. Ayrıca RF ile performans karşılaştırması sağlamak için NB, kNN ve SVM olarak üç sınıflayıcı daha tasarlanmıştır. Çekirdek parametreleri iyileştirmek amacıyla Grid arayış methodu ve SVM çekirdeği olarak radyal temel işlevi (RBF) kullanılmıştır. NB kullanılırken Gauss dağılım yöntemi seçilmiştir. kNN algoritması için, k değeri 3 ayarlanmıştır ve Öklid mesafesi kullanılmıştır. RF algoritması için ağaç sayısını ampirik olarak seçilmiştir. Değerlendirmelerde 10 kat çapraz doğrulama kullanılmıştır.

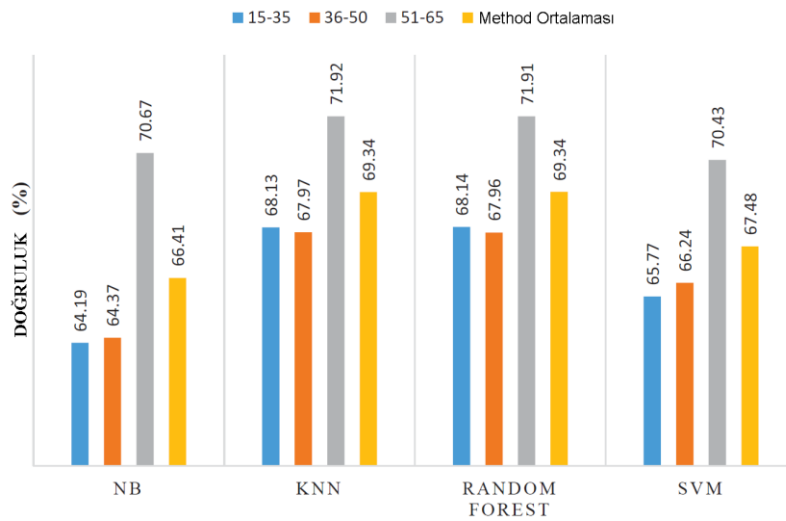
Oksijen düzeyindeki yaş faktörünü dikkate alarak yaş gruplarına dayanan üç test senaryosu hazırlanmıştır. Gruplara göre minimum, maksimum, ortalama ve standart sapma değerleri Çizelge 5.2'de verilmiştir. Standart sapma ile verilerin ortalamadan

ne kadar sapma gösterdiği izlenmiştir. 35-50 yaş grubunda veriler ortalamaya yakın dağılmıştır. Tersine, 51-65 yaş grubunda, veriler ortalamadan uzak yerlere dağılmıştır. Oksijen satürasyonu 51-65 yaş grubunda en düşük, en yüksek değerler ise tüm gruplarda aynıdır. Ortalama oksijen satürasyon değeri 51-65 yaş grubunda en düşüktür. Yani, Çizelge 5.2'deki istatistiklere göre, yaş arttıkça, kişinin kanındaki oksijen doygunluğunun azaldığı gözlenmiştir.

Çizelge 5.2 Farklı yaş grupları için istatistik değerleri

Yaş Grubu	Minimum	Maksimum	Ortalama	Standart Sapma
15-35	87	99	96,5	1,6
36-50	87	99	96,9	1,5
51-65	86	99	94,5	1,7

Oksimetre verilerini iç ve dış olmak üzere iki sınıfa ayrılmıştır. Şekil 5.2'de, üç farklı yaş grubu için test sonuçları sunulmuştur. Sonuçlara dayanarak, kişinin yaşı arttıkça, açık ve kapalı alan bulma sınıflandırma doğruluğu artmaktadır. Şekil 5.2'den okunabileceği gibi, oksimetre verisinden kişinin açık ya da kapalı alanda olup olmadığının ortalama tanıma doğruluğu 15-35, 36-50 ve 51-65 yaş grupları için sırasıyla %66.56 (64,19; 68,13; 68,14; 65,77), %66.63 (64,37; 67,97; 67,96; 66,24) ve %71.23 (70,67; 71,92; 71,91; 70,43) olarak belirlenmiştir. Sınıflandırma algoritmalarının doğruluğunu karşılaştırdığımızda, RF algoritmasıyla en yüksek % 69,34 ortalama doğruluk elde edilmiştir. kNN, SVM ve NB yöntemleri sırasıyla RF yöntemini takip eder.



Şekil 5.2 Farklı yaş grupları için kullanılan yöntemlerin, kişinin açık ya da kapalı alanda olup olmadığının tanıma doğruluğu.

Sonuç olarak, RF algoritması tüm senaryolarda en yüksek doğruluğu verir. kNN ve SVM algoritmalarının sonuçları da RF ile karşılaştırılabilir; Ancak, eğitim süresi yüksektir. Bu nedenle, bu yöntemler büyük veriler için tercih edilmeyebilir.

#### **5.5.1.2. Senaryo Bazlı Sınıflandırma Bulguları**

FA tanıma özelliğine sahip bir sistemi gerçekleştirmek için bir yaklaşım geliştirmek ve kullanmak önemlidir. Bu sistemle, akıllı telefonlar aracılığıyla elde edilen verileri kullanarak insan FA'sını doğru şekilde tahmin etmek mümkündür. Bu nedenle amaç, bir akıllı telefonda elde edilen verilerle statik durumu (ayakta durmak, oturmak) ve dinamik hareketi (yürümek, koşmak) başarıyla sınıflandırmaktır.

Tek bir sinyal modalitesi dikkate alınarak ya da aynı yaş grubundaki denekler üzerinde testler yapılmıştır (diğer bir deyişle 18-34 yaş insan sensör verisi, 35-55 yaş insan sensör verisi, 55-65 yaş insan sensör verisi). Bu nedenle, yaklaşımımızın etkinliğini göstermek için farklı yaş grubuna sahip farklı denekleri kullanan bir test senaryosu da tanımlanmıştır. Bu nedenle, fiziksel aktivite bulma methodumuzun başarımını bulmak için 3 tane senaryo hazırlanmıştır: Bu senaryolar, aynı kişi ile aynı yaş grubundakiler (WSwSA), farklı kişi ile aynı yaş grubundakiler (BSwSA) ve farklı kişi ile farklı yaş grubundakiler (BSwDA) arasındadır. Adlarından da anlaşılacağı gibi, WSwSA'nın eğitim ve test prosedürleri sırasında aynı yaş grubuyla aynı kişilerin örnekleri kullanılmıştır, yani, eğitim ve test aşamalarında A01 nesnesinin örnekleri kullanılmıştır. BSwSA'da, eğitim ve test sırasında farklı kişilerin ancak aynı yaş grubundaki örnekleri kullanılır. Yani, A01'i test etmek için A02, A03 ve A04 kişileri seçilmiştir. Son olarak, bu çalışma için yeni olan BSwDA ve algoritmaların sağlamlığını değerlendirmek için, eğitim ve test prosedürlerinde farklı yaş gruplarına sahip farklı kişilerin örnekleri kullanılmıştır. Bu durumda, A01'i test etmek için, tüm yaş grubundaki (A02, A03, A04; B01; B02; B03; B04; B04; C01, C02, C03; C04) diğer kişiler kullanılmıştır. BSwSA ve BSwDA senaryolarında, eğer eğitimde bir kişiyi kullandıysak testlerde asla aynı kişiyi kullanılmamıştır.

Bu çalışmada dört fiziksel aktivite tanımlanmış ve aşağıdaki gibidir:

- Ayakta durma
- Oturma
- Yürüyüş
- Koşma

Bu dört aktivite, bir kişinin genellikle yaptığı temel statik ve dinamik faaliyetlerdir. Bunlar, FA'yı başarıyla tanımada önemli bir rol oynar. Ayrıca, bu faaliyetler hem iç hem de dış mekânda iki farklı yerde gerçekleştirilmiştir. Bu görüşe göre deneklerin yaptıkları aktiviteyi nerede yaptıklarını bulmaya çalışılmıştır. Veriler, kullanıcının cep sağ pantolon cebinde bulunan telefonda ve sağ ellerinde bulunan telefonda toplanmıştır. Bu iki yer akıllı telefon ile FA tanıma için başarılı yerlerdir [86].

Testler, 4 farklı makine öğrenme yöntemi kullanılarak ACC, CLI, MAG ve GYR sensör verileri üzerinde yapılmıştır. Ayrıca performansları karşılaştırılmıştır. Yöntemlere göre bulunan fiziksel aktivite sınıflandırmasının doğruluk oranları Çizelge 5.3'de gösterilmiştir.

Çizelge 5.3 Fiziksel aktivite sınıflandırma metotlarının karşılaştırması

	WSwSA (%)				BSwSA (%)				BSwDA (%)			
	NB	kNN	SVM	RF	NB	kNN	SVM	RF	NB	kNN	SVM	RF
Çok Kipli Hepsi	94,68	99,82	<b>100</b>	<b>100</b>	70,43	51,42	47,84	69,7	74,46	55,73	56,02	78,94
Çok Kipli El	89,74	98,61	98,56	<b>99,99</b>	56,95	40,05	56,69	46,48	61,62	45,47	44,98	61,23
Çok Kipli Pantolon	89,05	97,4	97,96	<b>99,99</b>	64,35	50,74	51,91	55,85	70,86	62,7	56,96	67,33
ACC El	44,26	50,03	51,26	<b>78,55</b>	39,54	37,82	41,71	40,45	39,48	37,36	<b>43,24</b>	40,24
GYR El	34,75	42,17	32,72	<b>69,81</b>	30,53	30,75	<b>33,27</b>	30,99	31,7	30,65	<b>34,74</b>	31,26
MGN El	44,08	90,46	96,02	<b>90,9</b>	19,16	17,2	<b>20,31</b>	17,69	22,22	18,36	<b>24,21</b>	19,68
CLN El	48,45	77,71	98,02	<b>83,41</b>	23,26	20,47	<b>25,46</b>	21,13	25,38	21,81	<b>26,71</b>	23,04
ACC Pantolon	48,74	65,59	61,71	<b>86,31</b>	37,93	36,57	38	38,03	37,24	38,62	38,44	<b>40,39</b>
GYR Pantolon	33,6	41,51	33,77	<b>45,17</b>	29,77	31,83	29,58	<b>31,92</b>	29,87	<b>31</b>	30,65	<b>31</b>
MGN Pantolon	57,52	93,43	95,93	<b>97,82</b>	29,88	25,9	<b>31,31</b>	25,71	<b>31,43</b>	27,46	27,45	26,52
CLN Pantolon	61,19	81,4	99,83	<b>90,91</b>	27,55	26,26	<b>31,88</b>	27,84	<b>32,59</b>	29,28	31,84	29,12

### 5.5.2. MHealth Veri Kümesi Çalışmaları

Kullandığımız veri kümesi füzyon ve öznitelik seçme methodlarının başarımını tesbit etmek için verilerin tek ve çok kipli durumlarına göre ayarlanmıştır. Bu ayarlama amacıyla makine öğrenme methodlarımızın başarımını bulmak için değişik senaryolar oluşturulmuştur. Bu oluşturulan senaryolar yardımıyla verilerin değerler olarak birbirleri arasındaki mesafe yakınlığı olan verilerin etiketlerinin farklılık durumuna bakmadan verilerin ağırlığına bakarak seçme işlemi uygulanmıştır. Bununla beraber, sınıflandırma etiket değerinin bulunduğu özniteliklerin bir kısmı, bunu barındırmıyorsa, bu özniteliklerin makine öğrenme algoritmasına verilmesi anlamsız olacağını görmemiz sebebiyle, makine öğrenme algoritması



çalışmasından önce etiket durumunu içerisinde bulunduran ve makine öğrenme algoritmasının çalışmasını kolaylaştıran verileri bulmak için öznitelik seçme methodu kullanılmıştır [102]. Bu methoda göre bulunan özniteliklerin makine öğrenme algoritması başarımına katkıları Çizelge 5.4'de sunulmuştur. Bu değerlere göre çok kipli şeklinde füzyon sırasında, sırasıyla minimum katkı barındıran öznitelik çıkarılarak işlem yapılmıştır.

Çizelge 5.4 Verilerin tek kipli şeklinde fiziksel aktivite tahminindeki başarımına etkisi (0: en düşük, 1: en yüksek)

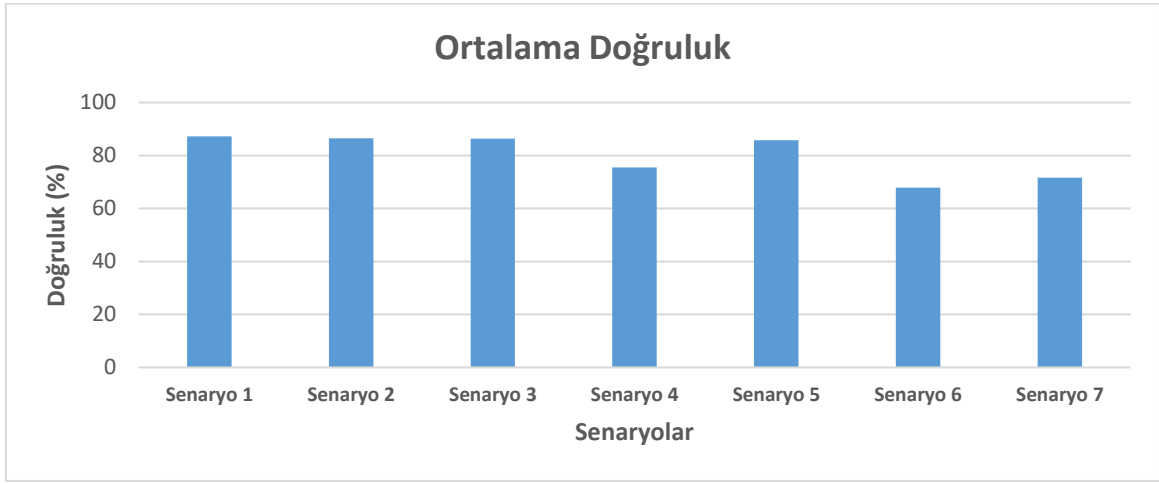
<b>Sensör Adı</b>	<b>Relieff</b>	<b>Bilgi Kazancı</b>
<b>GYR</b>	0,100	0,550
<b>ACC</b>	0,050	0,370
<b>MAG</b>	0,020	0,320
<b>EKG</b>	0,009	0,110

Tek ve çok kipli durumunda değişik senaryolar olarak makine öğrenme yöntemlerini kullandığımızda elimizdeki senaryolar olarak en yüksek başarıım değeri çok kipli olarak RBF SVM ile oluştuğu bulunmuştur. Farklı senaryoların ve makine öğrenme yöntemlerinin gösterildiği değerler Çizelge 5.5'de gösterilmiştir.

Çizelge 5.5 Senaryo (Sen) bazlı fiziksel aktivite sınıflandırması için makine öğrenme yöntemlerinin doğruluklarının karşılaştırma tablosu

<b>Sen. Numarası</b>	<b>RBF DVM (%)</b>	<b>Poly. DVM (%)</b>	<b>NB (%)</b>	<b>kNN (%)</b>	<b>RF (%)</b>	<b>Bagging (%)</b>
Sen. 1	<b>98,40</b>	81,57	75,02	94,82	87,75	85,78
Sen. 2	<b>94,85</b>	81,13	74,26	94,74	87,80	85,87
Sen. 3	<b>93,72</b>	79,43	79,01	92,89	87,78	85,43
Sen. 4	77,22	72,86	66,75	73,88	<b>81,90</b>	80,12
Sen. 5	<b>96,27</b>	76,76	76,30	95,11	85,83	84,54
Sen. 6	70,73	71,24	46,52	70,99	<b>74,54</b>	73,37
Sen. 7	71,23	71,72	67,86	71,34	<b>74,51</b>	73,34

Değerleri incelediğimizde, çekirdek olarak RBF kullandığımızda ve özneliğin seçimi yapmadan Senaryo 1 için çok kipli senaryo ile %98,40 makine öğrenme başarımlarını aldığımız görülmektedir. Senaryo 5 olan GYR verisi sıradaki en yüksek değeri (%96,27) veriyorken; Senaryo 2 için öznelik alma methoduna göre etkisi minimum olan EKG çıkarıldığı çok kipli yöntem ile %94,85 başarımlar üçüncü sırada gelmektedir. Senaryolar göz önüne alınarak oluşturulan ortalama başarımlar Şekil 5.3'de gösterilmektedir.



Şekil 5.3 Senaryo bazlı fiziksel aktivite sınıflandırması için makine öğrenme yöntemlerinin ortalama doğrulukları

İlk senaryo olan senaryo 1 ile değerleri incelediğimizde çekirdek olarak RBF kullanılan DVM, %98,40 başarımlarıyla en yüksek değeri vermektedir. DVM methodunu kNN %94,82 ikinci sırada bulunmaktadır. Bir başka makine öğrenme algoritmalarından olan ve yüksek doğruluğu veren rastgele ağaç methodunun başarımları %87,75'dir. kNN ile rastgele ağaç değerleri arasında başarımlar farkı fazladır. İlgili senaryo kullanıldığında minimum başarımları NB makine öğrenme yöntemi seçildiğinde %75,02 ile olduğu bulunmuştur. Yapılan çalışmalara göre, çekirdek olarak RBF seçilmenin DVM başarımlarını artırmıştır. Ayrıca başka makine öğrenme yöntemlerinin çok kipli kullanımının sınıflandırma başarımlarını azalttığı sonucunu çıkarmaktadır.

Tek kipli veri analizi kullanıldığında maksimum başarımların %96,27 ile senaryo 5 yani GYR sensörü kullanıldığında bulunan makine öğrenme algoritması ile bulunduğu görülmüştür. Bu bulunan sonuç, senaryo 2 ve senaryo 3'deki çok kipli yapılan

çalışma sonucunda bulunan sonuçlardan da fazladır. Bulunan bu değerler, öznelik seçme methodları yardımıyla bulunan GYR'nin sınıflandırma etiketini bilgi olarak maksimum bulunduran öznelik olması durumunu göstermektedir. Diğer modların bulunduğu senaryo 4, senaryo 6 ve senaryo 7 başarımlarını incelediğimizde, çok kipli senaryolar için diğer topluluk yöntemlerinden Rassal Ağaç yöntemi ile Bagging Tree yönteminin doğruluğu yüksektir. Bu senaryolar kullanıldığında, diğer sırada çekirdek olarak RBF kullanılan DVM gelmektedir.

Dahası olarak, çok ve tek kipli makine öğrenme algoritmalarımızın sonuçları geçmişteki benzer çalışmalar ile karşılaştırılarak sonuçlar tartışılmıştır. [88 - 89]'de gösterilen çalışmalar, kullandıkları veri kümesi aynı olmasından dolayı ve çok kipli olmalarından dolayı seçilmiştir. Chanyaswad vd.'nin önerdikleri method SNR tabanlıdır [88]. Önerilen yöntemimiz olan veri füzyon ve veri çıkarma methodu ile değişik sayılarda veriler kullanıldığında başarımlarımız [88]'e göre %11,07 yüksektir. Diğer çalışma olan Kutlay vd.[89] testleri iki farklı sınıflandırma yöntemi içermektedir ve eğitim işlemi aşamasında değişik veri kümeleri kullanmaktadırlar. Kullandığımız yöntemler ile karşılaştırılan çalışmaların karşılaştırılması Çizelge 5.6'da sunulmuştur. Kutlay vd. en yüksek başarımları %91,70 ile Yapay Sinir Ağları ile almışlardır. Çekirdek fonksiyonu RBF olan çok kipli SVM yöntemimiz, öznelik seçerek %98,40'lık başarımlar sağlamaktadır. Yapılan çalışmalar göstermiştir ki, veri füzyonu ve öznelik çıkarma methodu ile oluşturulan verileri kullanmak makine öğrenme algoritması doğruluğunu yükseltmektedir.

Çizelge 5.6 Kullanılan method ile [88] ve [89] yöntemleri arasındaki başarımların karşılaştırılması

Method	RBF çekirdek DVM (%)	SNR tabanlı makine öğrenme (%)	Polinomial çekirdek DVM (%)	Yapay Sinir Ağları(%)
Yaklaşımımız	96,270	-	81,570	-
Yaklaşımımız (Öznelik Seçimi kullanmadan)	<b>98,400</b>	-	-	-
Chanyaswad vd. [88]	-	87,330	-	-
Kutlay vd. [89]	-	-	<b>83,220</b>	<b>91,700</b>

### 5.5.3. Physionet Veri Kümesi Çalışmaları

Tek kipli çalışmalarda, OUA sorununun bulunması için tek bir sensör verisi kullanılmaktadır. EKG ve SpO<sub>2</sub> sinyallerinden toplanan iki bağımsız sensör verisi

kullanılmıştır ve bunların öznitelik düzeyindeki füzyon işlemi gerçekleştirilmiştir. EKG ve SpO<sub>2</sub> sinyallerine dayanan öznitelik seviyesi füzyonunun amacı, sınıflandırıcının eğitiminden önce her sinyal tarafından yakalanan çeşitli ayrıntıları kullanmaktır. Birleştirme, toplam havuzlama ve maksimum havuzlama gibi farklı kaynakların özelliklerini füzyon etmek amacıyla değişik methodlar kullanılabilir. Ayrıca, toplam ve maksimum havuzlama methodları, verilerin barındırdığı önemli bilgileri kaybedebilirler. Bundan dolayı, EKG ve SpO<sub>2</sub> verileri ile bulunan önemli bilgiyi saklamak için füzyon methodu kullanılmıştır. 5.2 bölümünde ayrıntılı olarak anlatılan füzyon yöntemi kullanılmıştır.

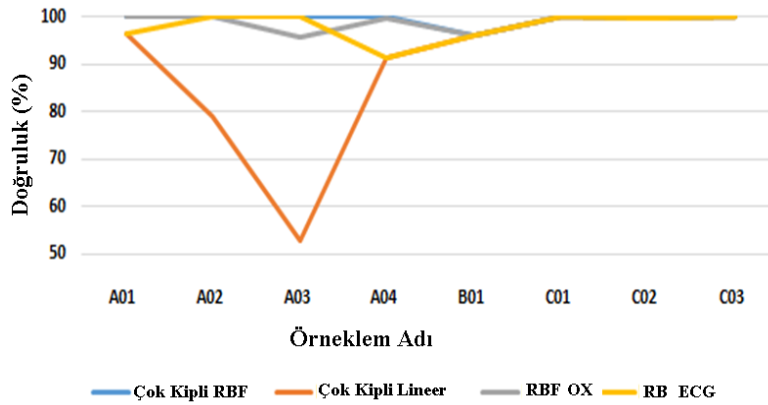
Tek bir sinyal modalitesini göz önünde bulundurarak veya aynı zamanda uyku apnesinin şiddeti ile ilgili (örneğin yüksek derecede apne, düşük derecede apne veya apnesi olmayan) testler de yapıldığı için, diğer mevcut yöntemlerle karşılaştırıldığında, farklı kullanılan bir test senaryosu da tanımlanmıştır. Bundan dolayı, makine öğrenme methodlarımızın başarımını karşılaştırmak amacıyla 3 değişik senaryo hazırlanmıştır: Bunlar Aynı Şiddete Sahip Kişiler (WSwSS), Aynı Şiddete Sahip Farklı Kişiler (BSwSS) ve Farklı Şiddete Sahip Farklı Kişiler (BSwDS) şeklindedir. Adlarından da anlaşılacağı gibi, WSwSS için eğitim ve test prosedürleri ve test aşamalarında aynı şiddete sahip aynı kişilerin örnekleri kullanılmıştır. BSwSS'de, eğitim ve test sırasında farklı kişilerin örnekleri fakat aynı şiddete sahip örnekleri kullanılır. Yani, A01, A02, A03 ve A04 kişilerinin tüm örnekleri kullanılır. Son olarak, bu çalışma için yeni olan ve algoritmaların sağlamlığını değerlendirmek için BSwDS'de, eğitim ve test prosedürlerinde farklı uyku apnesi şiddeti olan farklı kişilerin örneklerini kullanmayı öneririz. Bu durumda, testlerde farklı şiddete sahip deneklerin tüm örnekleri (A01, 02, 03, 04; B01; C 01, 02, 03) kullanılmaktadır. Tüm senaryolarda, eğer eğitimde kullandıysak, testlerde asla bir aynı kişiyi kullanmayız.

İlk test SVM yöntemi kullanılarak yapılmıştır. Buradaki amacımız hangi çapraz doğrulama (cross-validation) değerinin en başarılı olduğunu bulmaktır. Farklı sayıda CV'nin üzerindeki etkisi değerlendirilmiştir [104]. Bu amaçla, RBF tabanlı SVM sınıflandırıcı ile  $k = 2, 5, 7, 10$  seçilmiştir ve çok kipli sonuçları Çizelge 5.7'de sunulmaktadır. Sonuçlar, tüm senaryolar arasında  $k = 10$  olduğunda çok kipli yaklaşım için en iyi sonuçları elde ettiğimizi göstermektedir. Bu nedenle,  $k = 10$  seçilmiştir ve SVM'yi Linear çekirdeği kullanarak aynı test tekrar edilmiştir.

Çizelge 5.7 Çapraz doğrulama fold (katlama [kat]) karşılaştırması

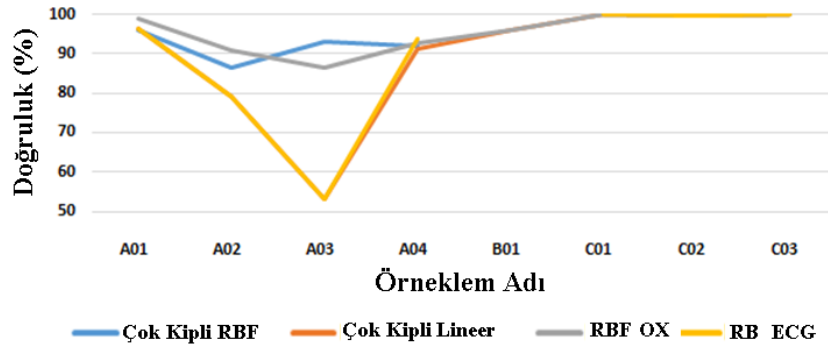
	WSwSS (%)				BSwSS (%)				BSwDS (%)			
	2 Kat.	5 Kat.	7 Kat.	10 Kat.	2 Kat.	5 Kat.	7 Kat.	10 Kat.	2 Kat.	5 Kat.	7 Kat.	10 Kat.
A01	100	100	100	100	95,69	95,88	95,91	96,11	94,07	94,19	94,3	94,53
A02	100	100	100	100	84,55	86,47	86,01	86,6	95,42	95,37	95,34	95,31
A03	100	100	100	100	92,03	92,49	93,01	93,07	95,78	95,84	96,87	95,81
A04	100	100	100	100	91,55	91,22	91,63	91,82	96,33	96,39	96,39	96,33
B01	96,09	96,09	96,09	96,09	-	-	-	-	95,28	95,37	95,4	95,57
C01	100	100	100	100	100	100	100	100	94,05	94,2	94,34	94,71
C02	99,8	99,8	99,8	99,8	99,8	99,8	99,8	99,8	93,59	93,65	93,73	93,82
C03	100	100	100	100	100	100	100	100	94,04	94,11	94,278	94,41

Ayrıca veri kümesindeki bireysel fizyolojik sinyallerin performansını değerlendirmek için tek kipli analizler yapılmıştır. Bu amaçla, bağımsız sinyal özelliklerini kullanarak her senaryo için (WSwSS, BSwSS ve BSwDS) üç test daha yapılmıştır. Her bir testte EKG, SpO<sub>2</sub> ve her ikisine birden dayanan SVM sınıflandırıcılı Doğrusal ve RBF bazlı çekirdekleri dikkate alırız. Şekil 5.4'de WSwSS senaryosu için test sonuçları gösterilmiştir. Testler 10 kat çapraz doğrulama (CV) kullanarak yapılmıştır. Sonuçlar, belirli bir test senaryosunu gerçekleştirdikten sonra tüm kişilerin ortalama doğruluğu olarak verilmiştir. Sonuçlar, RBF çekirdeğini kullandığınızda çok kipli yaklaşımla %99.49 en iyi sınıflandırma doğruluğunu elde ettiğimizi göstermektedir. SpO<sub>2</sub> sinyali ikinci en iyi sonucu verir (% 98.91) ve EKG %97.92 doğrulukla üçüncü sıradadır. Sonuçlar ayrıca, RBF çekirdeğinin tüm sinyaller için Doğrusal çekirdekten daha üstün olduğunu da göstermektedir.

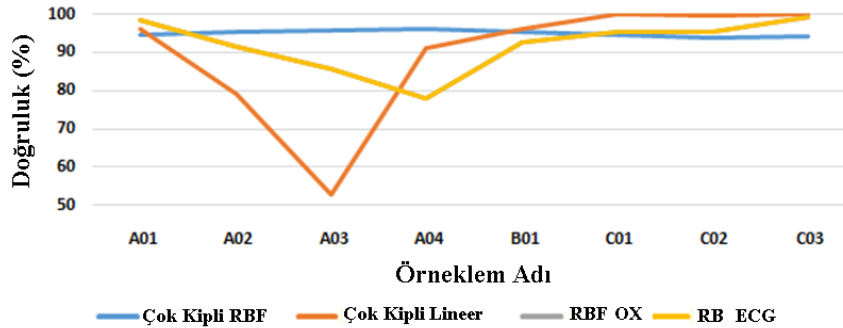


Şekil 5.4 WSwSS test senaryosu için tek ve çok kipli özneliklerin uyku apnesi sınıflandırması için performans karşılaştırması

BSwSS olan ikinci senaryo için, önceki test ile aynı test yapılmıştır ve Şekil 5.5'deki sonuçlar bulunmuştur. Bu senaryoda, SpO<sub>2</sub> RBF çekirdeğini kullanarak % 95,60 ile en iyi doğruluğu verir ve çok kipli yaklaşımla rekabet eder (% 95,34 ). EKG sinyali bu senaryoda % 89,43 (Doğrusal çekirdek) doğrulukla üçüncü sırayı alır. Sonuçlara dayanarak, SpO<sub>2</sub> sinyalinin EKG sinyalinden ziyade apnedeki değişikliklerini yakalamada daha iyi olduğunu söyleyebiliriz ve dolayısıyla EKG sinyali süper öznelik vektörüne (fs) hükmeder. Şekil 5.6'de son senaryonun (BSwDS) sonuçları sunulmuştur. Burada, RBF tabanlı SVM ile çok kipli yaklaşım, bağımsız sinyal öznelikleri arasında % 95.07 hassasiyetle en iyi performansı vermektedir. SpO<sub>2</sub> ve EKG sinyalleri ikinci sırayı % 92.03 doğrulukla paylaşıyor. Grafikten de okunabileceği gibi, çok kipli yaklaşım, tüm sinyaller için bağımsız sinyal özellikleriyle karşılaştırıldığında üstün tanıma oranlarına sahiptir. Çok kipli yaklaşımla yapılan genel iyileşme, kullanılan veri kümesi için % 3,04'tür.



Şekil 5.5 BSwSS test senaryosu için tek ve çok kipli özneliklerin uyku apnesi sınıflandırması için performans karşılaştırması



Şekil 5.6 BSwDS test senaryosu için tek ve çok kipli özneliklerin uyku apnesi sınıflandırması için performans karşılaştırması

Uyku apnesi sınıflandırma görevi için NB, kNN ve SVM algoritmaları arasında performans karşılaştırma yapılmıştır. Üç test senaryosuyla birlikte hem tek hem de çok kipli durumlar ele alınmıştır (Çizelge 5.8). SVM için Doğrusal ve RBF çekirdeklerini göz önünde bulundurur ve deneylerde kNN algoritması için  $k = 3$  ayarını yaparız. Çok kipli durumda, RBF tabanlı SVM algoritması, sırasıyla % 99.49, % 95.35 ve % 95.37 hassasiyetle WSwSS, BSwSS ve BSwDS senaryoları için en iyi sonuçları verir. Sonuçlar açıkça göstermektedir ki, tek kipli EKG özelliğinin sınıflandırma doğruluğu, WSwSS senaryosuna kıyasla BSwSS ve BSwDS senaryoları için çarpıcı biçimde azaldığını göstermektedir. Bu sonuç, EKG sinyalinin, çok kipli öznitelikler kadar farklı nesnelere özelliklerini yakalamada iyi olmadığını gösterir. Son olarak, yalnızca SpO<sub>2</sub> özelliğini kullanırken üç algoritmanın performansını karşılaştırırız. Yine SVM, NB ve SpO<sub>2</sub> özniteliği için kNN algoritmalarına kıyasla en iyi performans gösteren algoritmadır. Çok kipli öznitelikler ile karşılaştırıldığında, kazanan çok kipli özelliğidir. Ayrıca, bağımsız SpO<sub>2</sub> sinyali EKG sinyalinden daha üstündür.

Ayrıca, en iyi performans gösteren çok kipli RBF tabanlı SVM sınıflandırıcımızın sonuçlarını literatürdeki ilgili çalışmalarla karşılaştırılmıştır. Literatürde başka çalışmalar olmasına rağmen, aynı performans veri kümesini, PhysioNet'i kullandıkları ve çok kipli yaklaşımları göz önüne aldıkları için [96 - 97] 'deki çalışmalar seçilmiştir. [96] 'de Shi vd. "Apnea MedAssist II" adı verilen veri kümesi kullanmışlardır. İlgili veri kümesi OUA takip işlemi için SpO<sub>2</sub>, EKG ve nefes sensörlerinden gelen verileri kullanır. Physionet veri kümesindeki sonuçları Çizelge 5.9'da sunulmaktadır. Tablodan okunabileceği gibi, bizim yöntemimiz WSwSS ve BSwSS test senaryoları için [96] ile karşılaştırıldığında daha iyi sonuçlar vermektedir (ortalama olarak % 3,87 daha yüksek). BSwDS test senaryosu bu çalışma için yeni olduğu için, [96] için geçerli değildir. Diğer karşılaştırma için kullanılan çalışma için, Xie vd. EKG ve SpO<sub>2</sub> verilerinin bağımsız ve kombine formlarda doygunluğuna bakarak online uyku apnesi sınıflandırması yapmışlardır [97]. Denemelerinde on makine öğrenme algoritması vardır ve öğrenme için farklı özellik kümeleri seçer. Bu yöntemlerin sonuçları yaklaşımımızla birlikte Çizelge 5.10'da gösterilmektedir.

Çizelge 5.8 NB, SVM ve kNN modellerinin uyku apnesi sınıflandırması için performans karşılaştırması

<b>Çok Kipli</b>			
	<b>WSwSS(%)</b>	<b>BSwSS</b>	<b>BSwDS</b>
Naive Bayes	92,19	88,86	89,13
Lineer Çekirdek SVM	89,43	89,43	89,43
RBF Çekirdek SVM	<b>99,49</b>	<b>95,35</b>	<b>95,07</b>
K En Yakın Komşu	97,65	80,96	75,21
<b>Tek Kipli (EKG)</b>			
	<b>WSwSS(%)</b>	<b>BSwSS</b>	<b>BSwDS</b>
Naive Bayes	99,64	88,4	85,10
Lineer Çekirdek SVM	<b>100,00</b>	89,43	<b>89,43</b>
RBF Çekirdek SVM	97,92	<b>88,85</b>	88,86
K En Yakın Komşu	<b>100,00</b>	88,18	89,38
<b>Tek Kipli (OX)</b>			
	<b>WSwSS(%)</b>	<b>BSwSS</b>	<b>BSwDS</b>
Naive Bayes	87,09	85,17	81,91
Lineer Çekirdek SVM	99,49	93,62	91,45
RBF Çekirdek SVM	98,91	<b>95,60</b>	<b>92,03</b>
K En Yakın Komşu	<b>100,00</b>	93,52	90,95

Çizelge 5.9 Metodumuzun [96] 'da olan çalışmalarla doğruluk karşılaştırması

	<b>WSwSS (%)</b>	<b>BSwSS (%)</b>	<b>BSwDS (%)</b>
Önerilen Yöntem	<b>99,490</b>	<b>95,350</b>	<b>95,080</b>
[96]	97,300	89,800	Mevcut Değil

Çizelge 5.10 Metodumuzun [97] 'da olan çalışmalarla doğruluk karşılaştırması

	<b>Doğruluk(%)</b>	
Önerilen Yöntem	<b>96,520</b>	
[97]	SVM	82,240
	C4.5	80,910
	Bagging Reptree	84,400
	Bagging ADtree	79,850
	FT Trees	79,320
	AdaBoost	77,790
	REP Tree	81,330
	K En Yakın Komşu	81,650
	Karar Ağaçları	80,790
	Çok Katmanlı Sinir Ağı	81,600



3.2.1. bölümünde ayrıntılı olarak anlatılan Relieff yöntemi, sınıflandırma problemlerinde en sık kullanılan yöntemlerden birisidir. Relieff methodu çift sınıf barındıran makine öğrenme yöntemlerinde yüksek doğruluklar sunmaktadır. Bundan dolayı, makine öğrenme yöntemlerimizden önce öznitelik seçimi olarak relieff seçilmiştir. Kullanılan veri kümesinde EKG ve SpO<sub>2</sub> verileri için ilk olarak öznitelik seçim işleminden sonra füzyon edilen verilerimiz için de ayrıca öznitelik seçim işlemi uygulanmıştır. 50 ile 2500 öznitelik arasında sınıflandırma algoritmasının etiketlerine göre verilerin uzaklıklarına bakılarak öznitelikler seçilmiştir [103]. Daha sonra da sonuçlar karşılaştırılmıştır. Örn. İlk 500 yüksek öncelikli öznitelik bulmak amacıyla ilgili seçme methodundan bulunan ağırlıklara göre en yüksek 500'ü kullanılmıştır. İlgili öznitelik seçme algoritması ile bulunan özniteliklere bağlı makine öğrenme algoritması başarımları Çizelge 5.11'de sunulmaktadır. En fazla başarımlar 1000 ve 2000 değerleri için bulunmuştur. Burada bulunan en fazla doğruluk değeri olan 1000 makine öğrenme algoritması eğitimlerinde kullanılmıştır.

Çizelge 5.11 Kullanılan öznitelik seçme methodu ile bulunan özniteliklerin sayılarına göre uyku apnesi sınıflandırma için makine öğrenme başarımları (\*: Öznitelik seçimi yapılmadan)

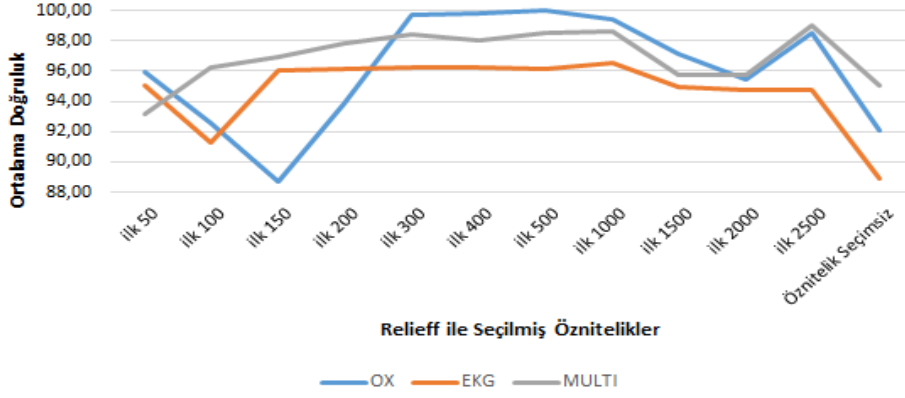
Kullanılan Öznitelik Adedi		Füzyon Edilmiş Öznitelikler	Doğruluk (%)
EKG	SpO <sub>2</sub>		
6000* adet	6000* adet	12000 adet	91,99
50 adet	50 adet	100 adet	94,71
100 adet	100 adet	200 adet	93,37
150 adet	150 adet	300 adet	93,88
200 adet	200 adet	400 adet	95,94
300 adet	300 adet	600 adet	<b>98,12</b>
400 adet	400 adet	800 adet	<b>98,02</b>
500 adet	500 adet	1000 adet	<b>98,22</b>
1000 adet	1000 adet	2000 adet	<b>98,21</b>
1500 adet	1500 adet	3000 adet	95,94
2000 adet	2000 adet	4000 adet	95,31
2500 adet	2500 adet	5000 adet	97,43

Kullandığımız sensör verisi füzyon ve öznitelik seçimi yöntemlerinin başarımlarını test etmek amacıyla aynı ve farklı OUA şiddetine göre kişiler (yani yüksek apne, düşük apne derecesi ya da apnesiz) ayrılmıştır. Ayrıca verilerin tek ve çok kipli çeşitleri de bu aşamada düşünülmüştür. Bunlardan dolayı makine öğrenme algoritmamızın

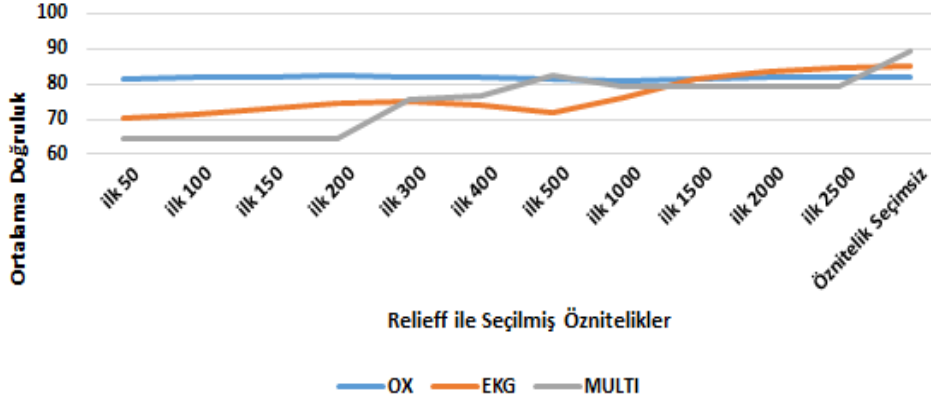
başarımını bulmak amacıyla değişik senaryolar oluşturulmuştur. Oluşturulan senaryolara göre öznitelikler değer açısından birbirlerine benzer olan verilerin etiketlerinin aynı ya da farklı durumuna göre özniteliklerin ağırlıkları değiştirilmiştir.

Tek ve çok kipli testlerde makine öğrenme algoritmalarını denediğimizde kullanılan senaryoya göre en yüksek başarımlar çok kipli analiz ile RBF çekirdeğini kullanan SVM ile bulunmuştur. Sonuçlar incelendiğinde, çekirdek olarak RBF seçtiğimizde, ayrıca öznitelik seçimi işleminden sonra çok kipli olarak %97.13 makine öğrenme algoritması başarımlarını bulmaktadır. SpO<sub>2</sub> verisi sıradaki en yüksek doğruluğu (%96,45) vermektedir. EKG ile yapılan çalışmalarda da %95,28 başarımlarını sıradaki en yüksek olur. Relief methodundan bulunan sonuçlar sıralandığında da en yüksek 50 ile en yüksek 2500 öznitelik seçilmektedir. Bu aşamada SpO<sub>2</sub>, EKG ve ikisinin füzyon edilmiş versiyonu olan çok kipli verilerden ayrıca öznitelik seçme işlemi de yapılmıştır. Bu işlemden sonra da daha önce en yüksek başarımlar bulduğumuz çekirdek seçimi olarak RBF olan SVM makine öğrenme methodu yeniden yapılmıştır. SVM ile bulunan seçilmiş özniteliklerin başarımlarının karşılaştırması Şekil 5.7'de görülmektedir. Sonuçlara baktığımızda öznitelik sayısı düşük olduğunda sınıflandırma algoritmasının başarımları da düşük çıkmaktadır. Öznitelik sayısı büyüdükünde de sınıflandırma algoritmasının başarımları artmaktadır; bununla beraber 1000 adet öznitelik seçimi devamında öznitelik sayısı arttırılmaya devam ettiğinde başarımlarımız azalmaktadır.

NB için başarımlarını incelediğimizde SpO<sub>2</sub>, %81,77 ile en yüksek başarımlarını vermektedir ve çok kipli yaklaşım ile bulunan sonuçlarla (%73.67) arasında fark yüksektir. Bu durumda EKG verisi %75,97 başarımla sonuçlanmıştır. Sonuçları incelediğimizde, Naive Bayes makine öğrenme methodu kullanılarak SpO<sub>2</sub> verisinin kişilerin değişikliklerini EKG verisinden daha başarılı bir şekilde bulmaktadır. Bu nedenle, EKG verisinin uyku apnesi için etiketleri daha kötü temsil ettiğini söyleyebiliriz. Bununla beraber öznitelik seçimi uygulanmadığında başarımların SpO<sub>2</sub> sensörü haricinde EKG ve çok kipli olarak daha başarılı sonuçlar verdiği görülmüştür. Bu sonuçları incelediğimizde de, NB algoritması ile öznitelik seçimi uygulanacaksa SpO<sub>2</sub> sensörü kullanılabilir. EKG sensörü ile çok kipli sensör verisi kullanıldığında, makine öğrenme algoritmasının öznitelik seçimi ile başarımları düştüğü görülmektedir. Naive Bayes algoritması ile yapılan seçilmiş veriler ile başarımların sonuçlarının karşılaştırmalı grafiği Şekil 5.8'de görülmektedir.

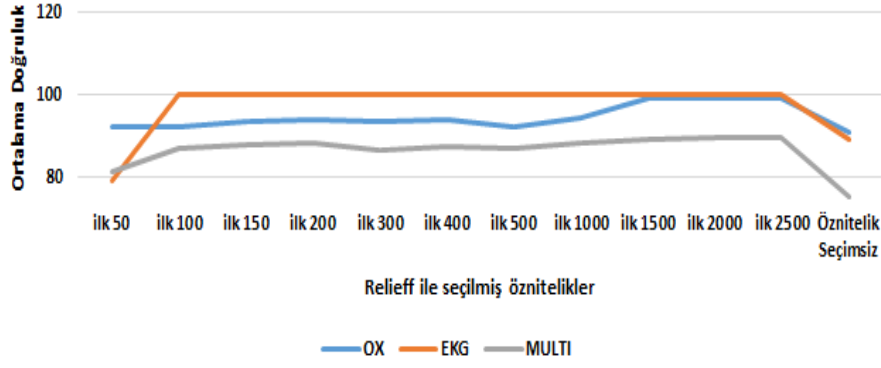


Şekil 5.7 Relieff ile seçilen öznitelikler kullanıldığında uyku apnesi için DVM doğruluk performansı



Şekil 5.8 Relieff ile seçilen öznitelikler kullanıldığında uyku apnesi için NB doğruluk performansı

K En Yakın Komşu algoritması için başarımları incelediğimizde EKG, %98,13 ile en yüksek başarımları vermektedir ve SpO<sub>2</sub> verisi (%94,9) ile başarımları farkı fazladır. Bu karşılaştırmada, çok kipli veriler seçildiğinde %87,56 başarımları üçüncüdür. Bunlardan dolayı, NB makine öğrenme methodu seçildiğinde EKG verisinin kişilerin farklılaştığı bilgisini SpO<sub>2</sub> verisinden daha iyi bulduğu görülmektedir. Bu nedenle, SpO<sub>2</sub> verisinin sınıf etiketini bulma aşamasında daha başarısız olduğunu sunabiliriz. Bununla beraber, öznitelik seçimi yapılmadığında başarımları SpO<sub>2</sub>, EKG ve çok kipli analizlerde düşük olduğu görülmektedir. Bu sonuçlara baktığımızda da k En Yakın Komşu methodu ile öznitelik seçimi uygulamak sınıflandırma başarımlarını arttırmak için önemlidir. k En Yakın Komşu methodu ile yapılan seçilen veriler ile doğruluğun başarımlarının karşılaştırmalı grafiğini Şekil 5.9'da görülmektedir.



Şekil 5.9 Relieff ile seçilen öznitelikler kullanıldığında uyku apnesi için kNN doğruluk performansı (k=3)

Bu işlemlerden sonra, çok ve tek kipli makine öğrenme algoritmalarımızın başarımları geçmişteki benzer çalışmalar ile karşılaştırılmıştır. Geçmişte farklı çalışmalar da olmasına rağmen, [96-97]'de bulunanlar, bizimle aynı veritabanı olmaları ve çok kipli methodları kullandığımız için seçilmiştir. [96] Shi vd. "Apnea MedAssist II" adı verilen veri kümesi kullanmışlardır. İlgili veri kümesi OUA takip işlemi için SpO<sub>2</sub>, EKG ve nefes sensörlerinden gelen verileri kullanır. Kullandığımız method veri füzyon ve öznitelik çıkarma methodu ile değişik sayıda veriler sayesinde doğruluğu [96]'e oranla %4.67 yükseltmiştir. Diğer çalışmada Xie vd. EKG ve SpO<sub>2</sub> verisini kullanarak online uyku apnesi sınıflandırması yapmışlardır[97]. Çalışmalarında farklı sayıda sınıflandırma methodu kullanılmıştır. Bu methodlar ile eğitim esnasında farklı sayıda veri kümesi seçilmiştir. Sonuçlara baktığımızda, %84.40 başarımla en yüksek başarımları Bagging Reptree vermiştir. Kullandığımız yöntemler ile literatürdeki yöntemlerin karşılaştırması Çizelge 5.12'de verilmiştir. Çekirdek fonksiyonu olarak RBF kullandığımız SVM, öznitelik seçimi işlemi ile %98.22'lük bir başarımla sağlamaktadır. Bu yüksek doğruluğun sağlanmasında veri füzyon ve öznitelik çıkarma methodlarını kullanmanın faydası bulunmaktadır.

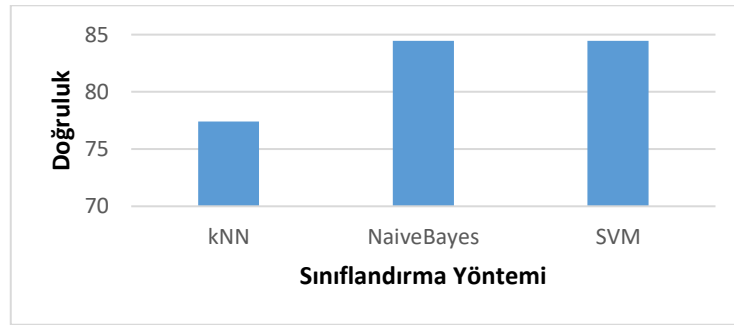
Çizelge 5.12 Kullanılan yöntemimiz ile [96] ve [97] çalışmalarında kullanılan en yüksek doğruluğu veren methodların karşılaştırması.

	DVM (%)	NB (%)	kNN (%)	Bagging Reptree (%)
Kullanılan Yöntem (Öznitelik Seçimi uygulanarak)	<b>98,22</b>	82,04	<b>96,30</b>	-
Kullanılan Yöntem (Öznitelik Seçimi uygulanmadan)	91,99	85,38	85,18	-
Shi vd. [96]	93,55	-	-	-
Xie vd. [97]	82,24		81,65	84,40

#### 5.5.4. Statlog Veri Kümesi Çalışmaları

Bu veri kümesinde kalp atım, yaş, EKG, tansiyon değişkenlerin Pearson Çarpım Moment yöntemi ile ilintilerini analiz ettiğimizde; kalp atım hızının yaş ile ilintisinin %40,22 çıktığını görmekteyiz. EKG verisi ile kalp atım hızı ve tansiyon değerleri arasındaki ilintinin çok düşük (%7 ve %11) değerlerde olduğu görülmüştür. Ayrıca, yaş ile kalp atım hızı arasında da %40 oranında ilinti olduğu görülmüştür.

Bu veri kümesinde kNN, Naive Bayes ve SVM sınıflandırma algoritmalarını uyguladığımızda; en başarılı sonucu Naive Bayes algoritması ile aldığımızı görmekteyiz. Farklı çapraz doğrulama yöntemleri ile test ettiğimiz veri kümesinde 5 fold yapılan Naive Bayes'in %85,19 doğruluk çıktığını görmekteyiz. Daha sonra SVM algoritması ile yapılan 10 fold sınıflandırma ile %84,81 doğruluk çıktığını görmekteyiz. kNN algoritması ile sınıflandırma yaptığımızda en yüksek sonucun 3 fold ile %79,63 doğruluk çıktığını görmekteyiz. Farklı fold değerlerine göre yapılan çalışmada ortalama doğruluk olarak Naive Bayes ile SVM eşit çıkmaktadır. Şekil 5.10 kalp hastalıklarına göre farklı sınıflandırma algoritmalarının doğruluk değerlerinin karşılaştırmasını göstermektedir.



Şekil 5.10 Kalp hastalıkları için farklı sınıflandırma algoritmalarına göre doğruluk değerlerinin karşılaştırması

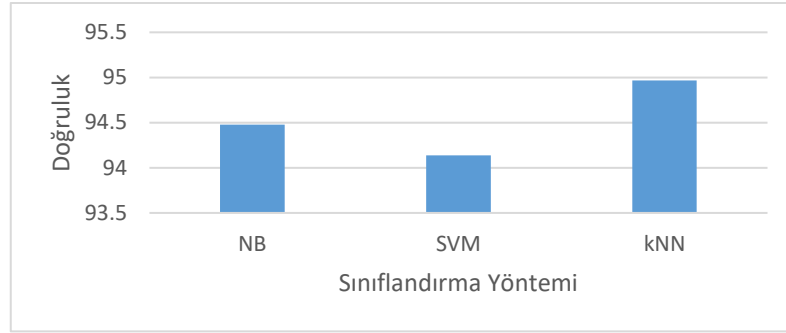
Aynı veri kümesi üzerinde bilgi kazancı yöntemi ile özneliklerin genel trendde ne kadar etki ettiğini bulmak istediğimiz çalışmada en yüksek bilgi kazancının göğüs ağrısı tipi ile kandaki floroskopi ile bakılan damar rengi çıkmıştır.

#### 5.5.5. MIT-BIH Veri Kümesi Çalışmaları

İlgili veri kümesi MLII ve V5 EKG sinyallerini içermektedir. Bu isimler standart EKG terminolojisini kullanan elektrot konumlarına karşılık gelir. MLII standart lead II'ye paralel iki kutuplu modifiye lead II'dir ve gövde üzerine yerleştirilen elektrotlar (uzun

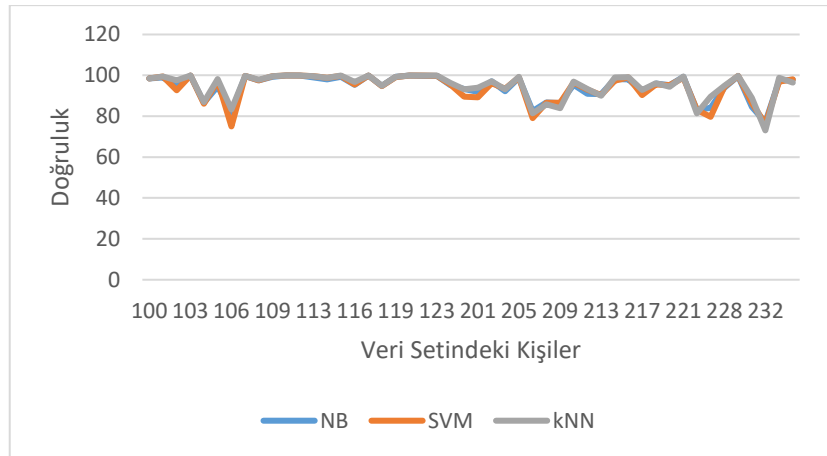
sürelili EKG monitörizasyonu için bir gereklilik) ile elde edilmiştir. V5, MLII'ye kabaca dik olan bir prekordiyal lead'tir. MLII'nin çoğu kişide yüksek genlikli normal QRS kompleksleri ürettiği ve V5'nin genellikle MLII'de düşük amplitüdde olan herhangi bir ektojik atımın hemen hemen en iyi frontal düzlem izdüşümünü sunduğu için bu iki veri birçok kayıt için tercih edilir.

Bu veri kümesinde kNN, Naive Bayes ve SVM sınıflandırma algoritmalarını uyguladığımızda; en başarılı sonucu kNN algoritması ile aldığımızı görmekteyiz. kNN ile yaklaşık %95 elde ettiğimiz ortalama doğruluk değeri NB ve VM ile %94 civarlarında çıkmıştır. Şekil 5.11 aritmi için farklı sınıflandırma algoritmalarına göre doğruluk değerlerinin karşılaştırmasını göstermektedir.



Şekil 5.11 Aritmi için sınıflandırma algoritmalarına göre doğruluk değerleri

Veri kümesinde bulunan 48 tane farklı kişinin verilerinin hangi noktada en düşük doğruluk verdiğine baktığımızda 106,207 ve 232 verilerinde en düşük değerleri verdiğini görmekteyiz (Şekil 5.12).



Şekil 5.12 Aritmi için sınıflandırma algoritmalarına kişi bazlı doğruluk değerleri

## 6. GÜRBÜZ ÖZNETELİK ÇIKARIMI YÖNTEMLERİ

Bu bölümde sensör verilerinden, sınıflandırma doğruluğunu artırmak için yeni öznetelik oluşturma işlemleri anlatılmaktadır. Bu anlatılan yöntemler sayesinde hem sınıflandırma doğruluğu artmış olacak hem de sınıflandırma için gerekli olan maliyet (zaman ve işlemci) azalmış olacaktır.

Birden fazla sensör kullanmak veri kümesindeki aykırı değerleri artırabilir; ancak, genel verileri tanımlamanın, her veri kümesinin özellikleri nedeniyle kesin bir yolu yoktur. Bunun yerine, ham gözlemleri yorumlamalı ve bir değer dışlayıcı olup olmadığına karar vermeliyiz.

Bununla beraber, zamansal olarak hareket eden sensör verilerindeki trendleri belirlemek ve sensörlerden gelen hareketin yönünü analiz etmek sınıfları tahmin etmede önemlidir. Çünkü sensör verisindeki eğilimi tahmin etmek ve bir eğilimin ne zaman tersine dönebileceğini bulmak ya da erkenden tahmin etmek, sınıfları bulmamızı daha da hızlandırabilir. Bu amaçlarla, yeni öznetelikler sunulmuştur.

### 6.1. Sigma Tabanlı Öznetelik Oluşturma

Sınıflandırma methodlarında eğitim verilerinin dağılımı sınıflandırma başarımına etkisi önemlidir. Eğitim verikümesinde uç veriler var ise, eğitim süresi uzayabilir ve sınıf etiketlerini öğrenme aşamasında problem olabilir. Bu problem de, sınıflandırma methodlarının eğitim sürecinde problemlere neden olabilir. Bu problemin üstesinden gelmek için verikümümüzdeki uç değerleri bulmak ve bu değerlerin değiştirilmesi amaçlanmıştır. Bu nedenle, sensör verilerinden yeni bir öznetelik üretilmiştir.

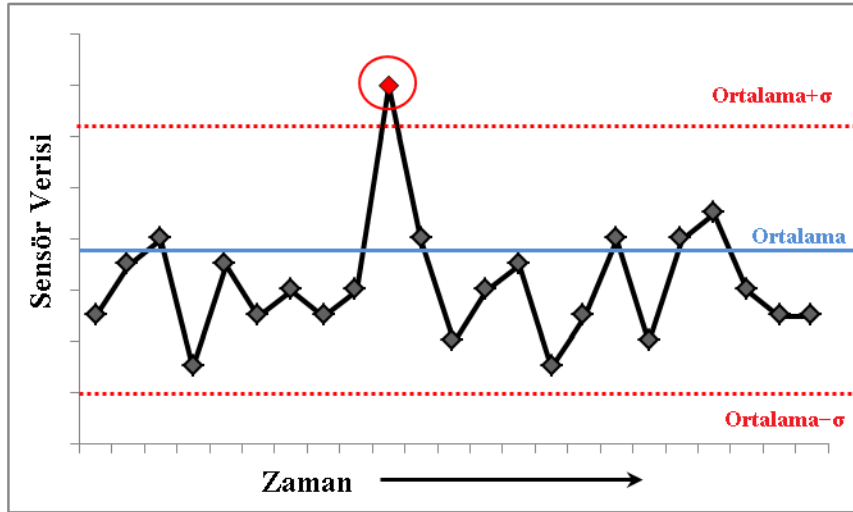
Bu motivasyonla, öncelikle veri kümesindeki verilerin ortalamaya göre dağılımı ölçülmüştür ve sigmanın ( $\sigma$ ) varyansın kare kökü olarak hesaplanmıştır (standart sapma). Hesaplama işlemi için (6.1) eşitliğindeki standart sapma hesaplama formülünü kullanılmıştır.

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2} \quad (6.1)$$

(6.1) eşitliğindeki  $x_i$ , bir sensör verisi örneğini ve  $\mu$  bu sensör verisinin ortalamasını belirtir.

Bu hesaplama bize, ortalamaya göre her veri noktası arasındaki değişimi bularak bu değeri varyansın karekökü olarak göstermektedir. Veriler ortalamasının uzağında

ise, veri kümesinde standart sapma yüksektir; bu nedenle, veri ne kadar geniş dağılımlı olursa, sapma değeri de yüksektir. Toplanan verikümesinde, her bir aktivite için 24 cep telefonu sensör sensör verisinden ilk olarak ortalama hesaplanmıştır. Daha sonra da ortalama standart sapma hesaplanmıştır. Bunun devamında sensöre ve aktiviteye göre en yüksek ve en düşük değerler ortalama  $\pm$  sigma ile hesaplanmıştır. Bunun devamında, uç değerlerin yerine başka değer atamak için ortalama  $\pm$  sigma değerinin dışındaki değerler bulunup, bunlar hesaplanan değerlerle değiştirilmiştir. Elimizdeki verilerden bir tanesi ortalama + sigma toplamından büyükse, bu veriler ortalama + sigma toplamı ile değiştirilmiştir. Bununla beraber, elimizdeki verilerden bir tanesi ortalama - sigma değerinden küçükse, bu veriler ortalama - sigma farkı ile değiştirilmiştir. Bu yöntemle, özneliliğin yeni gösterimi tamamlanmıştır ve sensör verilerindeki uç değerler temizlenmiştir. Şekil 6.1'de örnek ortalama  $\pm$  sigma gösterimi bulunmaktadır. Burada, grafikte gösterilen ortalama  $\pm$  sigma değerlerini aşan sensör verileri kırmızı kesikli çizgi ile gösterilen ortalama  $\pm$  sigma değeri ile değiştirilmiştir.



Şekil 6.1 Ortalama  $\pm$  Sigma Gösterimi Örneği

## 6.2. Üstel Hareketli Ortalama Tabanlı Öznitelik Oluşturma

Hareketli ortalamalar, bir sensör verisinin ortalama değerini belirli bir süre boyunca görselleştirir. Bununla birlikte, birkaç farklı hareketli ortalama türü vardır. Sensör verisindeki trendleri belirlemek ve sensörden gelen hareketin yönünü analiz etmek için hareketli ortalamaları (MA) kullanılmaktadır. Hareketli ortalamalar, sensör

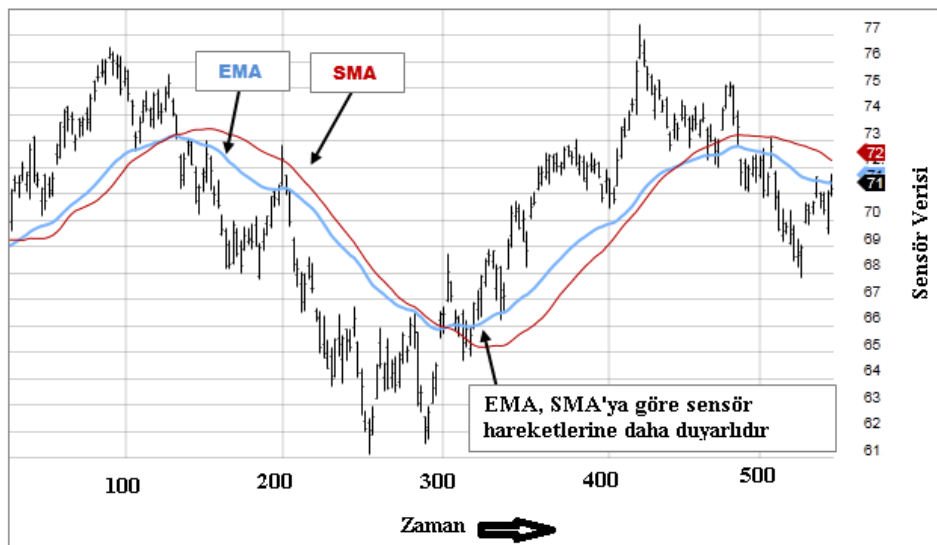


verisindeki eğilimi tahmin etmeye ve bir eğilimin ne zaman tersine dönebileceğini de işaret eder. En yaygın türlerden ikisi basit ve üsteldir.

Hareketli ortalamalar, sensör verisinin belirli bir süre boyunca ortalama değerini gösterir. Bununla birlikte, ortalamaları hesaplamanın farklı yolları vardır ve bu nedenle farklı hareketli ortalama türleri vardır. Bunlara "hareket etme" denir, çünkü sensör değeri değiştikçe hesaplamaya yeni veriler eklenir ve bu nedenle ortalamayı değiştirir.

Üstel Hareketli Ortalama (EMA) ve Basit Hareketli Ortalama (SMA), her birinin eğilimleri ölçmelerine benzerdir. İki ortalama da aynıdır çünkü aynı şekilde yorumlanırlar ve her ikisi de veri üzerindeki dalgalanmalarını düzeltmek için kullanılabilir. Bununla birlikte, iki ölçüm arasında bazı farklılıklar vardır. Bir EMA ile bir SMA arasındaki temel fark, her birinin hesaplamalarında kullanılan verilerdeki değişikliklere gösterdiği hassasiyettir. SMA, sensör verilerinin ortalamasını hesaplar, EMA mevcut verilere daha fazla ağırlık verir. En yeni verilerin, hareketli ortalamaları daha fazla etkileyecek ve eski sensör verilerinin etkisi daha düşük olacaktır. Daha spesifik olarak, üstel hareketli ortalama, son değerlere daha yüksek bir ağırlık verirken, basit hareketli ortalama tüm değerlere eşit ağırlık atar.

Bu nedenlerle, biz son gelen sensör verisinin, verinin trendi hakkında daha fazla bilgi vereceğini düşündüğümüz için üstel hareketli ortalamalar kullanılmıştır. Şekil 6.2'de SMA ile EMA karşılaştırması görülmektedir.



Şekil 6.2 Eğimölçer verisi kesiti için SMA ile EMA karşılaştırması

### 6.2.1. Tekil üstel hareketli ortalama (EMA)

Üstel Hareketli Ortalama (EMA) ayrıca Üstel Düzleştirme olarak da adlandırılır. EMA, hareketli ortalama tekniklerinin en iyisidir ve analizciler tarafından diğer hareketli ortalama yöntemlerine göre daha fazla tercih edilmektedir [80]. Davranışsal olarak, sensörlerden gelen yeni verilere yanıt olarak EMA, aşırı hassas ağırlıklı hareketli ortalama ile aşırı durgun basit hareketli ortalama arasında durur. Diğer ortalama tekniklerle karşılaştırıldığında, EMA mevcut verilerin eğilimini sorunsuz bir şekilde takip ederek atlamaları, kıpırdatmaları ve gecikmeleri en aza indirir.

Hesaplamalı olarak, EMA tüm hareketli ortalama tekniklerinin en basit ve en modern olanıdır [80]. EMA en az hesaplama, en az veri işleme ve en az veri geçmişi gerektirir. EMA sadece iki veri periyodu için sayısal değerler gerektirir. Bunlar en son elde edilen ham veriler ve bir önceki zamandaki EMA'dır. Örneğin, zamansal sensör verileriyle çalışarak, istenen bir zamanın EMA'sını hesaplamak için yalnızca istenilen zamanın gözlemlenmiş, işlenmemiş verilerine ve bir önceki zamanın EMA'sına ihtiyacımız vardır. Bu nedenle, EMA, uzun zamansal veri listelerini tutma ve yönetme ihtiyacını ortadan kaldırır.

EMA yönteminin önemli bir avantajı, EMA'nın eski veriler tarafından asla bozulmamasıdır. Eski veriler hiçbir zaman EMA değerini aniden düşürmez çünkü bu aslında hesaplamanın bir parçası değildir. Pratik amaçlar için, geçmiş EMA'nın giderek azalan ağırlığından dolayı geçmiş verilerin etkisi yavaş yavaş kaybolur. EMA'nın hesaplama yöntemi, eski verilerden gelen ve kullanılmayan verilerin hesaplamadan çıkarılmasından kaynaklanan düzensiz mevcut hareket sorununu doğru bir şekilde önler.

Tekli üstel Hareketli Ortalama aşağıdaki denklem ile hesaplanır:

$$EMA = (C - Ep) \times K + Ep \quad (6.2)$$

(6.2) formülünde; EMA, istenen zaman için Üstel Hareketli Ortalamayı; C, son zamandaki (hesaplanan zamandaki) sensör değerini; Ep, bir önceki dönem için Üstel Hareketli Ortalamayı; K üstel düzeltme sabitini ( $2 / (n + 1)$ )'e eşittir); n basit bir hareketli ortalamadaki toplam periyot sayısını göstermektedir. Üstel düzeltme sabit formülü  $K, 2 / (n + 1)$ , herhangi bir EMA'nın uzunluktaki n'nin daha yavaş Basit

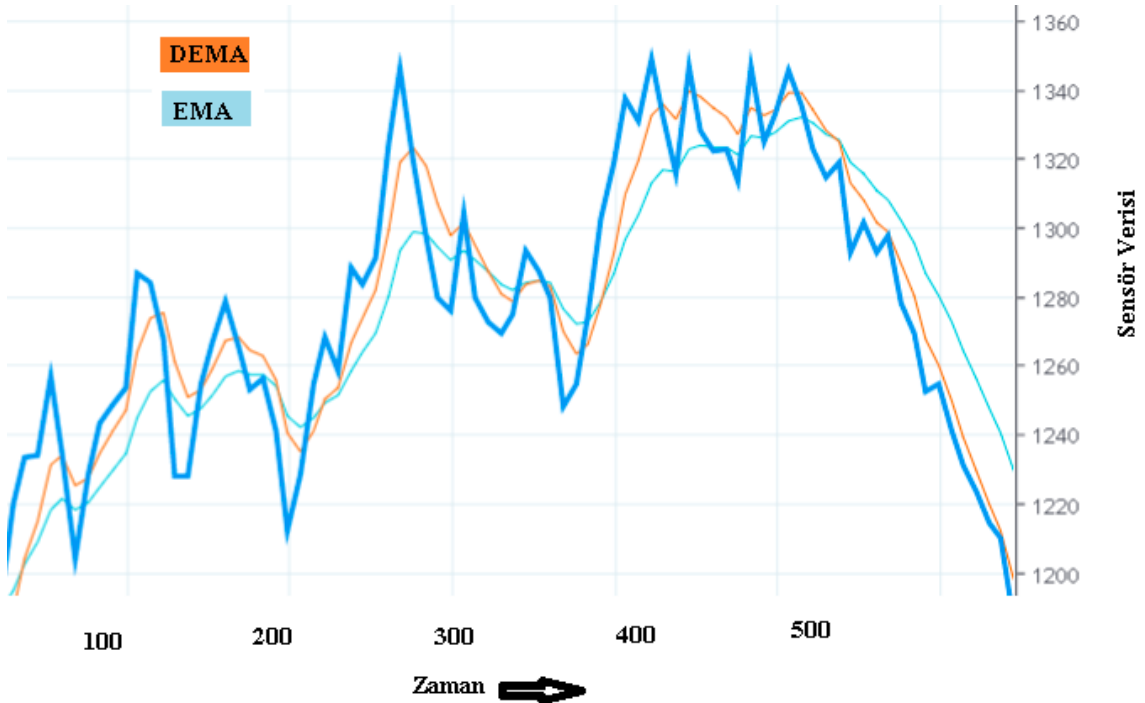
Hareketli Ortalama ile karşılaştırmasını sağlar.  $n$  sayısı arttıkça,  $K$  değeri gittikçe küçülür ve EMA yeni verilere karşı daha az hassas hale gelir.

### 6.2.2. Çiftli üstel hareketli ortalama (DEMA)

Çiftli Üstel hareketli ortalama (DEMA), basit üssel ve çiftli üssel hareketli ortalamaların bir kombinasyonudur. Bir başka deyişle DEMA hareketli ortalamasının hareketli ortalamasıdır. DEMA alışılmış üssel hareketli ortalamada meydana gelen gecikmenin azaltılması için tasarlanmıştır [80]. DEMA geleneksel hareketli ortalamalarda bulunan gecikme süresini azaltmak amacıyla geliştirilen daha yumuşak ve daha hızlı bir hareketli ortalama. Çiftli üstel hareketli ortalama (6.3) eşitliği ile hesaplanmaktadır.

$$DEMA = (2 * EMA(n)) - (EMA(EMA(n))) \quad (6.3)$$

DEMA'yı hesaplamanın ilk adımı EMA'yı hesaplamaktır. Ardından, ilk EMA hesaplamasının sonucunu kullanarak EMA hesaplaması tekrar çalıştırılmıştır. Son olarak, sonuç  $2 * EMA(n)$  değerinden çıkarılmıştır. Şekil 6.3'de EMA ile DEMA karşılaştırması görülmektedir. DEMA, EMA'ya göre sensör hareketlerine daha duyarlıdır.



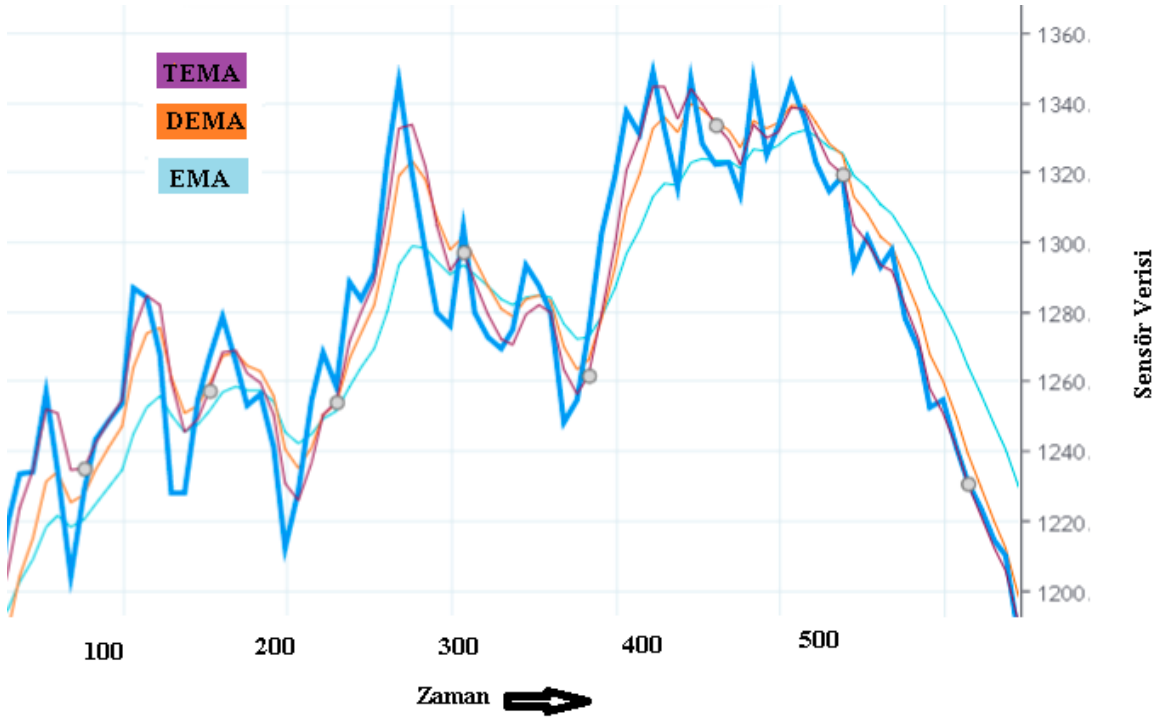
Şekil 6.3 Magnetometre verisi kesiti için DEMA ile EMA karşılaştırması

### 6.2.3. Üçlü üstel hareketli ortalama (TEMA)

Üçlü Üstel Hareketli Ortalamalar (TEMA), sinyalleri hızlandırmak ve sensör dalgalanmalarına daha hızlı yanıt vermek için üç farklı Üstel Hareketli Ortalama (EMA) kullanır [80]. TEMA, tek, çift ve üçlü EMA'ları kullanır. İlk EMA sensör değerini düzeltirken, ikinci EMA ilk EMA'yı, üçüncü EMA ise ikinci EMA'yı düzeltir. TEMA açısı, kısa süreli sensör verisinin yönünü belirtmek için kullanılabilir. TEMA formülü karmaşıktır ve gerçekte gecikmenin bir kısmını çıkarır. TEMA, bir sensör verisi artış trendini doğrulamaya yardımcı olur. TEMA, bir sensör verisi düşüş eğiliminin bulunmasına yardımcı olur. Üçlü üstel hareketli ortalama (6.4) eşitliği ile hesaplanmaktadır.

$$TEMA = (3 * EMA_1) - (3 * EMA_2) + EMA_3 \quad (6.4)$$

(6.4) formülünde  $EMA_1$ , sensör verisinin üstel hareketli ortalaması;  $EMA_2$ ,  $EMA_1$ 'in üstel hareketli ortalaması;  $EMA_3$  ise  $EMA_2$  değerlerinin üstel hareketli ortalamasıdır. Şekil 6.4'de TEMA ile EMA ve DEMA karşılaştırması görülmektedir. TEMA, DEMA ve EMA'ya göre sensör hareketlerine daha duyarlıdır.



Şekil 6.4 Magnetometre verisi kesiti için TEMA ile DEMA ve EMA karşılaştırması

## 7. SINIFLANDIRICI TASARIMI

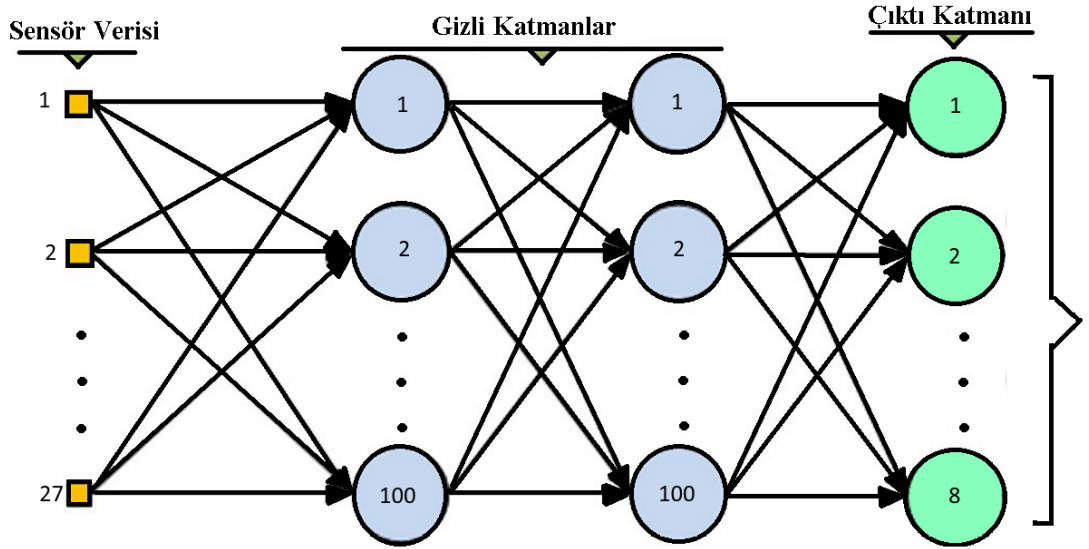
Makine öğrenmesi algoritmaları, akıllı telefonların veya giyilebilir sensörlerin yardımıyla insan hareketlerini sınıflandırmak için kullanılmıştır. Bu çalışmada kullanılan makine öğrenme yöntemleri alt başlıklar halinde anlatılacaktır. Bu çalışmada referans noktası oluşturmak için dört farklı sınıflandırıcı seçilmiştir. Bu sınıflandırma yaklaşımlarından dördü, ilgili çalışmalarda sıklıkla kullanılır ve çok çeşitli sınıflandırma karmaşıklığını temsil eder. Kalan dördü ise bizim tasarlamış olduğumuz derin öğrenme algoritmalarını içermektedir. Bu algoritmalar ile alakalı temel bilgiler 3. bölümde anlatılmıştır.

Derin öğrenme sınıflandırıcılarını kullanırken Deeplearning4j Java kütüphanesi kullanılmıştır. Deeplearning4j, Java ve Scala için yazılmış açık kaynaklı, dağıtılmış derin öğrenme kütüphanesidir. Hadoop ve Apache Spark ile entegre olan Deeplearning4j, kullandığımız sınıflandırma fonksiyonlarını dağıtılmış GPU'larda ve CPU'larda kullanılmak üzere iş ortamlarına getirir. Deeplearning4j, her biri katman olarak adlandırılan, basit yapay sinir ağlarından, derin sinir ağları oluşturmanıza olanak tanır. İçerisinde Derin öğrenen bir ağı eğitirken ayarlanması gereken birçok parametre bulundurmaktadır [98].

### 7.1. Derin Yapay Sinir Ağları (DNN)

Bu bölümde, bir kullanıcının FA'sını cep telefonundan ve oksimetre sensörü verilerinden öğrenen önerilen DNN yapısı tanıtılmaktadır. Bu çalışmanın arkasındaki motivasyon, cep telefonu ve oksimetre sensörlerinden gelen sıralı verileri kullanmak ve ayrıca gelişmiş bir FA sınıflandırması tasarlamak için DNN yapısını oluşturmaktır. Şekil 7.1, önerilen DNN yapısını göstermektedir. Sınıflandırma sırasında, bir DNN modelini eğitmek için kullanılan sensörlerden gelen veriler füzyon edilir.

Sensör verilerini kullanmak için, verileri füzyon etmemiz gerekir. Sensör sinyallerine dayalı öznitelik seviyesi füzyonunun amacı, sınıflandırıcının eğitiminden önce bu sinyal özelliklerini füzyon ederek her sinyal tarafından yakalanan çeşitli ayrıntıları kullanmaktır. Bu nedenle 5.2 bölümünde anlatılan öznitelik füzyonu işlemi uygulanmıştır.



Şekil 7.1 Derin Yapay Sinir Ağı Mimarisi

Bir DNN modeli, birden çok gizli katmandan oluşur. Bununla birlikte, aynı model ileri beslemeli yapay bir sinir ağı modelidir. Tüm gizli katmanlardaki nöronlar, bir sonraki gizli katmandaki nöronlara tamamen bağlanır. DNN modeli birden çok gizli katman içerdiğinden, giriş ve çıkış etiketleri arasında işlemleri verimli bir şekilde gerçekleştirebilir.

DNN modelinin  $L$  gizli katmanlardan oluştuğunu düşündüğümüzde, bu katmanın çıkışı (7.1) eşitliğindeki gibi tanımlanabilir:

$$a^l = f(w^l a^{l-1} + b^l) \quad (7.1)$$

burada  $l = 1, \dots, L$ ,  $a^0 = x$ , özellik vektörü  $x$  olan giriş katmanıdır.  $w^l$  ve  $b^l$ , nöral ağırlık matrisi ve  $l$ th gizli katmanın eğilimidir (bias).  $f(\cdot)$  doğrusal olmayan bir aktivasyon işlevidir [73]. Modelimizde Softmax aktivasyonu ile ileri beslemeli bir DNN tasarlanmıştır. Bu fonksiyon, softmax fonksiyonunun (7.2) eşitliğindeki gibi kabul edildiği ilgili gizli katmanların üstüne yerleştirilir.

$$\hat{y}_{i,j} = \frac{e^{a_j^l}}{\sum_d e^{a_d^l}} \quad (7.2)$$

Burada  $a_j^l$  ve  $a_d^l$  sırasıyla  $a^l$ 'nin,  $j$ th ve  $d$ th elemanlarıdır,  $\hat{y}_{i,j}$  ise  $\hat{y}_i$ 'nin  $j$ th elemanıdır ve ayrıca  $\hat{y}_i$ 'de  $x_i$  girdi verisinin  $i$ th elemanının çıktısıdır.

Stokastik Gradyan Azaltmasını (Stochastic Gradient Descent (SGD)) kullanıyoruz, böylece tüm veri kümesindeki her adımdan sonra ağırlık değerini güncelleyebiliriz.

Stokastik Gradyan Azaltmasında asıl amaç, veri kümesindeki her örnek üzerinde işlem yaptığınızda ağırlığı güncellemektir.

Sinir ağıın Xavier'in başlatılmasıyla eğitilmiştir [74]. Öğrenme hızını ampirik olarak seçeriz ve öğrenme oranının değerini kaybedilen işlevin azalmaya başladığı noktaya koyarız. Daha iyi ayarlanmış ağırlık güncellemeleri sağlamak için, öğrenme oranını 0,1 olarak seçeriz.

Parametre ayarlarını AdaGrad [75] optimizasyonu ile yapılmıştır. Bu aşamada, geri yayılım (backpropagation) algoritmasını kullanarak optimizasyon yapılmıştır. Geri yayılım sırasındaki ağırlık güncellemeleri (7.3) eşitliğindeki gibi yazılabilir:

$$w_{ij}(n + 1) = w_{ij}(n) + \mu \frac{\theta C}{\theta w_{ij}} \quad (7.3)$$

burada  $w$ ,  $C$  ve,  $\mu$  nöronun ağırlığını, (6.4) 'deki maliyet fonksiyonunu ( $C(Y, \hat{Y}, X, \theta)$ ) ve adım boyutunu sırasıyla belirtir [75]. Maliyet fonksiyonu (7.4) eşitliğindeki gibi gösteririz:

$$C(Y, \hat{Y}, X, \theta) = -\frac{1}{I} \sum_i \sum_j y_{i,j} \log(\hat{y}_{i,j}) \quad (7.4)$$

Burada  $\theta$  model parametrelerini gösterir.

Ampirik analizlere dayanarak mimari, her biri 100 nörondan oluşan iki gizli katmanla tasarlanmıştır.

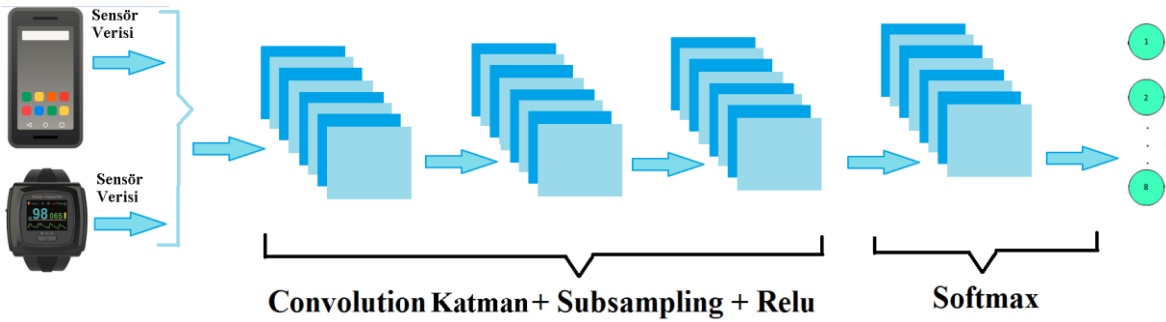
Deeplearning4j kütüphanesi kullanılarak oluşturulan mimarideki parametreler Çizelge 7.1'de gösterilmektedir.

Çizelge 7.1 Derin yapay sinir ağı için Deeplearnin4j kütüphanesi parametreleri

Parametre Adı	Değeri
İterasyon Sayısı	500
İlk Layer Aktivasyon Fonksiyonu	Activation.TANH
Başlangıç Ağırlık Fonksiyonu	WeightInit.XAVIER
Öğrenme Oranı	0,1
Regulasyon	TRUE
Ağırlıklar için L2 düzenleme katsayısı	e^-4
Gizli Katman Sayısı	2
Gizli Katmanlar için Nöron Sayıları	50
Çıktı Katmanı için Kayıp Fonksiyonu	LossFunction.NEGATIVELOGLIKELIHOOD
Çıktı Katmanı için Aktivasyon Fonksiyonu	Activation.SOFTMAX

## 7.2. Evrişimsel Sinir Ağı (ConvNet / CNN)

Elimizdeki topladığımız fiziksel aktivite verisini düşündüğümüzde, 27 öznitelikli fiziksel aktivite verisi bulunmaktadır. Bu veriyi CNN mimarisinde çalıştırabilmemiz için matris şekline dönüştürmemiz gerekmektedir. Bu dönüşüm sonunda 9x3 boyutunda bir veri kümemiz oluşur ve 3x3 boyutunda bir filtremiz bulunmaktadır. Filtre verimiz boyunca 1 adım sağ ve her yeni satırda bir adım aşağı kaydırılarak yeni veri kümesi elde edilir. Aktivasyon fonksiyonu olarak Relu kullanılmıştır. Mimarisinde 3 tane Convolution Layer, 3 tane Subsampling Layer, son olarak Softmax output layer kullanılmıştır. Filtre sayısı olarak farklı rakamları denemekle beraber en iyi sonucu veren 8 filtre sayısı kullanılmıştır. Subsampling temel olarak parametre sayısının düşürülmesi ve bakış açısındaki değişikliklerden ve uyarılardaki diğer bozulmalardan kaynaklanan bozulmalara tolerans getirilmesi için uygulanır. Biz her convolution layer'dan sonra subsampling uygulayarak parametre sayısı düşürülmüştür. Şekil 7.2 kullanmış olduğumuz CNN yapısının mimarisini göstermektedir. Deeplearning4j kütüphanesi kullanılarak oluşturulan Evrişimsel Sinir Ağı mimarisindeki parametreler Çizelge 7.2'de gösterilmektedir.



Şekil 7.2 Kullanmış olduğumuz CNN mimarisi

Çizelge 7.2 Evrişimsel sinir ağı için Deeplearnin4j kütüphanesi parametreleri

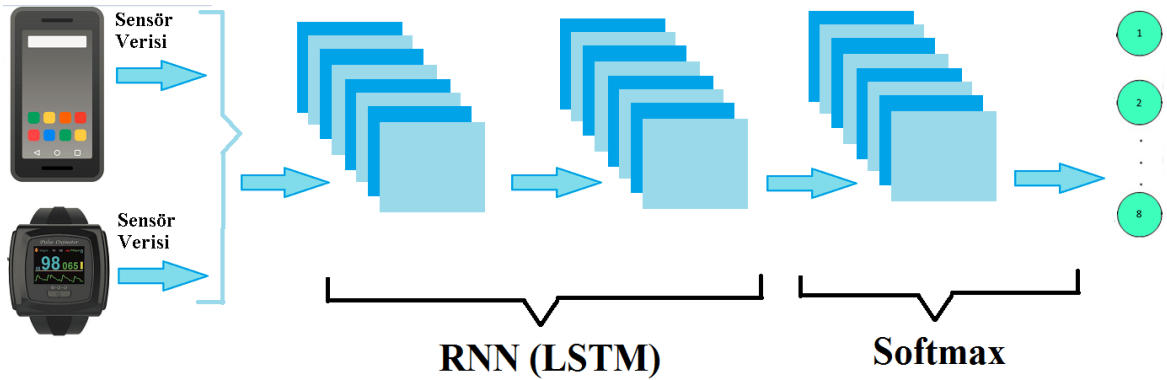
Parametre Adı	Değeri
Başlangıç Ağırlık Fonksiyonu	WeightInit.XAVIER
İterasyon Sayısı	500
Convolution Layer Aktivasyon Fonksiyonu	Activation.RELU
Filtre Matrisi Boyutu	3*3*1
Kerneldeki satır ve sütun sayısı	1
Subsampling eps değeri	e^-8
Subsampling havuz tipi	maksimum
Convolution Mode	Truncate
Çıkış Katmanı Aktivasyon Fonksiyonu	Activation.SOFTMAX
Çıkış Katmanı Kayıp Fonksiyonu	LossMCXent



### 7.3. Uzun / Kısa Süreli Bellek (Long / Short Term Memory (LSTM))

LSTM'lerin temelde RNN'lerden farklı bir mimarisi yoktur, ancak gizli durumu hesaplamak için farklı bir fonksiyon kullanırlar. LSTM'lerin hafızasına hücreler denir ve bunları önceki durum  $h_{t-1}$  ve geçerli giriş  $x_t$ 'yi girdi olarak alan kara kutular olarak düşünebilirsiniz. Dâhili olarak bu hücreler, ne tutacaklarına (ve ne sileceklerine) karar verirler. Daha sonra önceki durumu, mevcut belleği ve girdiyi birleştirirler. Bu tür birimlerin, uzun vadeli bağımlılıkları yakalamada çok etkili oldukları anlaşılıyor.

Bu çalışmada kullanılan RNN mimarisinde 2 tane LSTM layer kullanılmıştır ve arkasından Softmax output layer kullanılmıştır. Aktivasyon fonksiyonu olarak RELU fonksiyonu kullanılmıştır. Gate aktivasyon fonksiyonu olarak da RELU fonksiyonu kullanılmıştır. Şekil 7.3 kullanmış olduğumuz LSTM yapısının mimarisini göstermektedir. Deeplearning4j kütüphanesi kullanılarak oluşturulan Uzun/Kısa Süreli Bellek mimarisindeki parametreler Çizelge 7.3'de gösterilmektedir.



Şekil 7.3 Kullanmış olduğumuz LSTM mimarisi

Çizelge 7.3 Uzun/Kısa Süreli Bellek Sinir Ağı için Deeplearnin4j kütüphanesi parametreleri

Parametre Adı	Değeri
LSTM Katmanı Aktivasyon Fonksiyonu	Activation.RELU
Çıkış Katmanı Kayıp Fonksiyonu	Activation.SIGMOID
Çıkış Katmanı Aktivasyon Fonksiyonu	Activation.SOFTMAX

### 7.4. Hibrid CNN+LSTM Algoritması

Bu algoritma, esas olarak, CNN'ye dayanan özellik çıkarma aşamasına ve LSTM'ye dayanan füzyon aşamasına ayrılmıştır [79]. Özellik çıkarma aşamasında, sensör verilerinin ileri yayılma işlemi aşağıdaki gibidir: / katmanının bir evrimsel katman

olduğu ve  $l - 1$  katmanının bir havuz katmanı veya bir giriş katmanı olduğu varsayılmaktadır. Sonra  $l$  katmanının hesaplama formülü (7.5) eşitliğindeki gibidir.

$$x_j^l = f \left( \sum_{i \in M_j} x_i^{l-1} \times k_{ij}^l + b_j^l \right) \quad (7.5)$$

Yukarıdaki denklemin solundaki  $x_j^l$ ,  $l$  katmanının  $j^{\text{th}}$  öznitelik matrisini temsil eder. Sağ taraf,  $l - 1$  katmanın  $x_i^{l-1}$  ve  $l^{\text{th}}$  katmanın  $j_i$  konvolüsyon çekirdeği ile ilgili tüm özellik matrisi için evrişim çalışmasını ve toplamını gösterir. Daha sonra bir ofset parametresi ekler ve sonunda  $f (*)$  aktivasyon fonksiyonunu geçer. Bunların arasında,  $l$ , katman sayısıdır,  $f$ , etkinleştirme işlevidir,  $M_j$ , üst katmanın giriş öznitelik matrisidir,  $b$ , ofset ve  $k$ , evrişimli çekirdektir.

$l$  katmanın havuz katmanı (down sampling layer) olduğu varsayılarak,  $l - 1$  katmanı evrişimli katmandır.  $l$  tabakasının formülü (7.6) eşitliğindeki gibidir

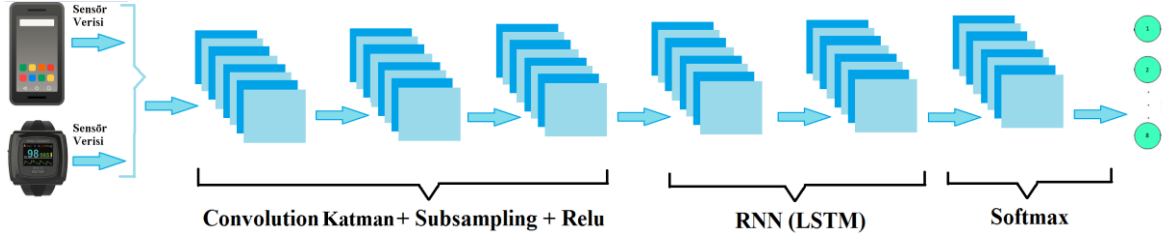
$$x_j^l = f \left( \beta_j^l \text{down}(x_j^{l-1}) + b_j^l \right) \quad (7.6)$$

(7.6) formülünde,  $l$  havuz katmanının sayısıdır,  $f$  aktivasyon işlevidir,  $\text{down} (*)$ , aşağı örnekleme işlevidir;  $\beta$  aşağı örnekleme katsayısı ve  $b$  ofset değeridir.

Sınıflandırma doğruluğumuzu arttırmak için kullandığımız CNN ve LSTM mimarilerini birleştirip yeni bir sınıflandırıcı yapmaktayız. Yaklaşımımız, literatürde en iyi performansı sağlayan CNN-RNN derin sinir mimarisine dayanmaktadır. Eğitim veri tabanımızı kullanarak eğitim prosedürünü başlatmak için belirli bir mimariyi kullanırız. Daha sonra, bu ağ yapısını, derin ağ yapısında çok sayıda bileşeni, ağ eğitiminde kullanılan hata ölçütlerini birleştirerek ve farklı işlem sonrası stratejileri uygulayarak genişletir ve uyarlarız. Geliştirilen yaklaşımların performansını validasyon veri kümemize göre değerlendirir ve elde edilen sonuçları sunarız.

Mimarimizde 3 tane Convolution Layer, 3 tane Subsampling Layer, 2 tane LSTM layer ve son olarak da Softmax output layer kullanılmıştır. Filtre sayısı olarak farklı rakamları denemekle beraber en iyi sonucu veren 8 filtre sayısı kullanılmıştır. Subsampling temel olarak parametre sayısının düşürülmesi ve bakış açısındaki değişikliklerden ve uyarılardaki diğer bozulmalardan kaynaklanan bozulmalara tolerans getirilmesi için uygulanır. Her convolution layer'dan sonra subsampling uygulayarak parametre sayısı düşürülmüştür. Aktivasyon fonksiyonu olarak RELU fonksiyonu kullanılmıştır. Şekil 7.4 kullanmış olduğumuz CNN + LSTM yapısının

mimarisini göstermektedir. Deeplearning4j kütüphanesi kullanılarak oluşturulan Hibrid CNN+LSTM mimarisindeki parametreler Çizelge 7.4'de gösterilmektedir.



Şekil 7.4 Kullanmış olduğumuz CNN + LSTM mimarisi

Çizelge 7.4 Hibrid CNN+LSTM Sinir Ağı için Deeplearning4j kütüphanesi parametreleri

Parametre Adı	Değeri
Başlangıç Ağırlık Fonksiyonu	WeightInit.XAVIER
İterasyon Sayısı	500
Convolution Layer Aktivasyon Fonksiyonu	Activation.RELU
Filtre Matrisi Boyutu	3*3*1
Kerneldeki satır ve sütun sayısı	1
Subsampling eps değeri	e^-8
Subsampling havuz tipi	maksimum
Convolution Mode	Truncate
Çıkış Katmanı Aktivasyon Fonksiyonu	Activation.SOFTMAX
Çıkış Katmanı Kayıp Fonksiyonu	LossMCXent
LSTM Katmanı Aktivasyon Fonksiyonu	Activation.RELU
Çıkış Katmanı Kayıp Fonksiyonu	Activation.SIGMOID
Çıkış Katmanı Aktivasyon Fonksiyonu	Activation.SOFTMAX

## 8. DENEYSEL BULGULAR

### 8.1. HPAwIO Veri Kümesi Bulguları

Hangi sensör verilerinin daha anlamlı olduğunu bulmak için sensör bazlı analizler yapılmıştır. Bu amaçla üç farklı çok kipli analiz ve sekiz farklı tek kipli analiz yapılmıştır. Çok kipli analizlerde, füzyon yöntemi kullanılmıştır. İlk olarak, tüm girdi verileri kullanılmıştır. Bunlar, hem el hem de pantolon cebinden toplanan dört cep telefonu sensörünün üç yönlü özelliğidir. Ayrıca, oksimetre, termometre ve kalp atış hızı sensörleri akıllı telefon sensörleri ile birlikte kullanılmıştır. İkinci çok kipli özneliklerimiz, pantolon cebindeki cep telefonu verilerinden, oksimetreden, termometreden ve kalp atışlarından oluşur. Son olarak, üçüncü çok kipli özelliğimiz, elde taşınan cep telefonu verilerinden, oksimetreden, termometreden ve kalp atış hızından alınan özelliklerden oluşur.

Testler ACC, CLI, MAG ve GYR sensör verilerinde 4 farklı makine öğrenme yöntemi kullanılarak yapılmıştır. Sensör verilerinin (özneliklerin) verimliliğini ölçmek için NB, KNN, SVM ve RF yöntemlerini temel yöntemler olarak kullanılmıştır. Ayrıca, 50 Hz civarında değerler FA tanıma için en yaygın kullanılan örnekleme oranlarından biridir [85]. Bu nedenle MATLAB kullanarak tüm veri kümesi 50 Hz örnekleme hızında yeniden örneklenmiştir ve daha sonra 50 Hz örnekleme oranına sahip tek ve çok kipli (füzyon edilmiş) sensörler için performans sonuçları sunulmuştur. Çizelge 8.1, çalışmamızda daha önce üzerinde çalıştığımız 200 Hz sonuçlarımızı göstermektedir. Bu tablo 50 Hz veri kümesi için de sonuçlarımızı da göstermektedir. Bu tabloda kullanılan tüm yöntemler aynıdır, tek fark hem 200 hem de 50 Hz veri kümelerinin kullanılmasıdır. Tabloların sonuçlarına göre, 200 Hz sonuçlarının 50 Hz sonuçlardan % 0-2 daha iyi olduğu görülmüştür ve veri kümemizdeki bulgularımız daha yüksek örnekleme oranlarının kullanılmasının tanıma oranını olumsuz yönde etkilemediğini doğrulamaktadır.

Deneylerimizin bir sonucu olarak, en iyi sonuç çok kipli veriler kullanıldığında RF yöntemiyle elde edilir. Tek sensörden gelen verilerin performansı değerlendirildiğinde, ACC sensöründen gelen verilerin en başarılı sonucu verdiği gözlemlenmektedir. Elden ve pantolon cebinden gelen verilerin sınıflandırma performansını kontrol ettiğimizde, sensörlerin pantolonun cebinden topladığı verilerin performansı üstündür. Çok Kipli Pantolondan elde edilen verilerin ortalama doğruluğu % 64,46 iken Çok Kipli Elden elde edilen verilerin ortalama doğruluğu %

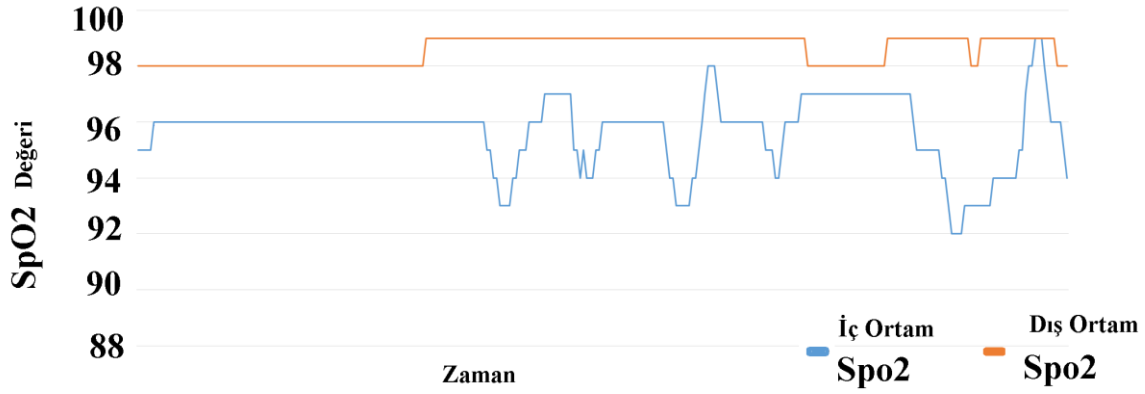
53,32'dir. Sonuç olarak, telefonun pantolon cebinde gerçekleştirdiği etkinliklerin daha net bir şekilde tanındığı gözlemlenmiştir. Bu sonuç, Kwapisz vd.'nin önerdiği gibi, vücutta sensörün doğru şekilde taşınmasının önemini doğrulamaktadır [86].

Çizelge 8.1 Tek kipli ile çok kipli veriler için örneklem oranlarına göre sensör FA sınıflandırma performansları (doğrulukları)

	200 Hz (%)				50 Hz (%)			
	NB	kNN	SVM	RF	NB	kNN	SVM	RF
Çok Kipli Hepsi	74,46	55,73	56,02	<b>78,94</b>	73,67	54,83	55,61	<b>78,60</b>
Çok Kipli El	<b>61,62</b>	45,47	44,98	61,23	<b>60,95</b>	44,68	44,86	60,84
Çok Kipli Pantolon	<b>70,86</b>	62,70	56,96	67,33	<b>70,41</b>	62,26	56,43	66,92
ACC El	39,48	37,36	<b>43,24</b>	40,24	38,83	36,93	<b>42,93</b>	39,60
GYR El	31,70	30,65	<b>34,74</b>	31,26	31,60	30,11	<b>33,87</b>	30,54
MGN El	22,22	18,36	<b>24,21</b>	19,68	21,37	17,59	<b>23,73</b>	19,30
CLN El	25,38	21,81	<b>26,71</b>	23,04	24,86	21,71	<b>26,58</b>	22,99
ACC Pantolon	37,24	38,62	38,44	<b>40,39</b>	36,69	37,83	37,67	<b>39,43</b>
GYR Pantolon	29,87	31,00	30,65	<b>31,00</b>	28,96	30,96	30,38	<b>30,88</b>
MGN Pantolon	<b>31,43</b>	27,46	27,45	26,52	<b>30,94</b>	27,02	26,99	25,97
CLN Pantolon	<b>32,59</b>	29,28	31,84	29,12	<b>32,17</b>	28,87	31,72	28,50

Farklı sensörlerden gelen verilerin ortalama doğruluk oranları incelendiğinde, ACC verilerinin sınıflandırılmasının en yüksek doğruluk sonucuna ulaştığı Çizelge 8.1'den görülmektedir. Pantolon cebinde ACC, MAG, CLI ve GYR verileri için ortalama doğrulukların sırasıyla % 38.67, % 28.21, % 30.71 ve % 30.63 olduğunu bulduk. Analizlerimize göre, MAG sensörünün ACC, GYR ve CLI sensörlerinden daha az belirgin olduğu açıktır. Sonuçlardan, ACC, GYR ve CLI verilerinin başarısının FA'ler için belirgin olduğu açıktır. Ayrıca, manyetik alanlar farklı FA'lar sırasında değişmektedir.

Öte yandan, FA'nın kapalı veya açık bir yerde yapılıp yapılmadığını çözmek için oksimetre sensörü kullanılmıştır. Açık ya da kapalı alanlar için farklı oksijen doygunluk seviyeleri beklenmektedir [101]. Bu amaçla FA'lar sırasında SpO<sub>2</sub> seviyelerinin grafiğinden anlaşılacağı üzere, SpO<sub>2</sub> seviyelerinin açık ve kapalı alanlarda nasıl farklılaştığı gözlemlenmiştir (Şekil 8.1).



Şekil 8.1 Açık ile kapalı alanlar arasında SpO<sub>2</sub> değişimi

SpO<sub>2</sub> sinyalinin etkisini daha iyi analiz etmek için, sadece SpO<sub>2</sub> sinyalini kullanarak açık ya da kapalı alan sınıflandırması yapılmıştır. Bazı veri kümesi istatistiklerinin yanı sıra sınıflandırma sonuçları Çizelge 8.2'de sunulmaktadır. Tabloda görüldüğü gibi, yaş arttıkça, bir kişinin kanındaki oksijen doygunluğu azalır. Benzer şekilde, kişinin yaşı arttıkça, açık ya da kapalı alan için sınıflandırma doğruluğu da artar. A, B ve C grupları için sırasıyla % 66,56, % 66,67 ve % 72,23 ortalama doğruluk oranları elde edilmiştir.

Çizelge 8.2 Oksimetre sensörü kullanma esnasında açık ya da kapalı alan sınıflandırma performansları (A: 18-34 yaş aralığı, B: 35-54 yaş aralığı, C: 56-65 yaş aralığı)

Veri Kümesi İstatistik Değerleri					Sınıflandırma Doğruluk Değerleri (%)			
Min	Maks	Ort.	Std. Sap.		NB	kNN	RF	SVM
A1	85	99	96,4	1,41	61,40	66,18	66,20	61,65
A2	90	99	96,81	1,53	63,16	64,11	64,16	63,37
A3	92	99	95,77	1,61	79,94	83,22	83,19	83,19
A4	82	99	97,08	1,97	52,28	59,00	59,02	54,88
B1	80	99	96,32	1,72	71,26	75,18	75,17	75,11
B2	91	99	97,26	1,12	60,18	60,17	60,18	60,18
B3	91	99	96,68	2,07	65,71	71,54	71,54	65,71
B4	87	99	97,44	1,29	60,31	64,97	64,94	63,95
C1	85	99	94,58	1,82	86,74	86,84	86,84	86,74
C2	88	99	95,41	1,51	57,92	62,71	62,70	57,93
C3	81	99	93,74	2,26	74,12	74,17	74,15	73,53
C4	89	98	94,25	1,25	63,88	63,95	63,96	63,51

Sensör deneylerine dayanarak, SpO<sub>2</sub> sensörü açık ve kapalı alanlarda beklendiği gibi etkisini açıkça gösterir ve bu nedenle çevreyi ayırt etmede iyi bir özneliktir. Ayrıca, sensör füzyonu (çok kipli) FA tanıma için bireysel sensör kullanımına kıyasla daha iyi doğruluklar sağlar ve çalışma için motivasyonumuzu oluşturur.

### 8.1.1. Senaryo bazlı sınıflandırma bulguları

Testler, rasgele ağaç makine öğrenme yöntemi ve 7.1 Bölüm'ünde tanıtılan DNN mimarisi kullanılarak ACC, CLI, MAG ve GYR sensör verileri üzerinde yapılmıştır. Ayrıca performansları karşılaştırılmıştır. Yöntemlere göre elde edilen doğruluk oranları Çizelge 8.3'de gösterilmektedir.

Çizelge 8.3 Açık ya da kapalı alan içeren FA için sınıflandırma methodlarını karşılaştırması

	WSwSA (%)		BSwSA (%)		BSwDA (%)	
	RF	DNN	RF	DNN	RF	DNN
Çok Kipli Hepsi	<b>100</b>	<b>100</b>	69,7	<b>70,91</b>	78,94	<b>79,04</b>
Çok Kipli El	<b>99,99</b>	97,34	46,48	<b>57,2</b>	61,23	<b>67,14</b>
Çok Kipli Pantolon	<b>99,99</b>	96,16	55,85	<b>67,88</b>	67,33	<b>72,73</b>
ACC El	<b>78,55</b>	48,11	40,45	<b>42,44</b>	40,24	41,68
GYR El	<b>69,81</b>	39,19	30,99	29,69	31,26	29,48
MGN El	<b>90,9</b>	74,9	17,69	16,84	19,68	17,83
CLN El	<b>83,41</b>	65,13	21,13	21,19	23,13	22,63
ACC Pantolon	<b>86,31</b>	50,75	38,03	<b>39,44</b>	<b>40,39</b>	39,47
GYR Pantolon	<b>45,17</b>	36,12	<b>31,92</b>	28,71	<b>31</b>	29,2
MGN Pantolon	<b>97,82</b>	77,79	25,71	23,17	26,52	25,64
CLN Pantolon	<b>90,91</b>	70,51	27,84	25,53	29,12	26,83

3 farklı çok kipli analiz ve 8 farklı tek kipli analiz yapılmıştır. Öncelikle tüm giriş verileri kullanılmıştır. Bunlar dört cep telefonu sensöründen gelen üç yönlü boyuttur. Ayrıca bu sensörler 2 farklı yerde kullanılmıştır. Ayrıca oksimetre, termometre ve kalp atış hızı kullanılmıştır. İkinci çok kipli veri, yalnızca pantolon cebinde, oksimetre, termometre ve kalp atış hızındaki cep telefonu verileridir. Üçüncü çok kipli veriler yalnızca elde tutulan cep telefonu verileri, oksimetre, termometre ve kalp atış hızıdır.

Yapılan deneyler sonucunda, en iyi sonuç çok kipli yaklaşımla önerilen DNN yöntemi ile elde edilmektedir. Tüm senaryolarda, önerilen çok kipli yaklaşım DNN yöntemi en başarılı sonuçları verir. Tek sensörlerden gelen verilerin performansı değerlendirildiğinde, MAG sensöründen gelen verilerin WSwSA senaryosunda en

başarılı sonuçları verdiği görülmektedir. Öte yandan, diğer senaryolarda ACC sensörü en başarılı sonucu verir. Tüm senaryolarda, tek kipli kullandığımız zaman, önerilen DNN yöntemimiz genellikle tatmin edici olmayan bir sonuç verir. Bu sonuçlar, DNN'de büyük bir veri kümesinin önemini doğrulamaktadır.

Elde tutulan telefonda ve pantolonun cebindeki telefonda gelen verinin sınıflandırma performansına baktığımızda; Pantolon cebinden gelen verilerden yapılan sınıflamanın ortalama doğruluğu daha yüksektir. Çok kipli pantolondan elde edilen verilerin ortalama doğruluğu %73,46 iken çok kipli elden elde edilen verilerdeki ortalama doğruluk %68,14'tür. Buradan çıkabilecek sonuçlara baktığımızda, pantolon cebindeki telefonla, etkinlikler daha belirgindir. Bu sonuç, Kwapisz vd.'nin [86] önerdiği gibi vücutta sensörlerin doğru şekilde yerleştirmenin önemini de doğrulamaktadır.

Farklı sensörlerden gelen verilerin ortalama doğruluk oranlarına baktığımızda; ACC verilerinin sınıflandırılmasının en yüksek doğruluk sonucuna ulaştığını görebiliriz. Pantolon cebinde ACC için % 46.48 doğruluk ortalamasını, pantolon cebinde manyetometre verileriyle sınıflandırma için % 46.46, pantolon cebinde GYR verileriyle sınıflandırma için % 32,91'dir. Çalışmamızda GYR sensöründen elde edilen verilerle yapılan sınıflamanın doğruluğu, ACC, MAG ve CLI sensörlerinden gözle görülür derecede düşüktü. Sebeplere baktığımızda, çalışmamızda sadece FA tanıma işlemi yapılmakla kalmıyor, aynı zamanda bir ortam içinde FA'yı da tanıma yapılmıştır. Diğer bir deyişle, açık ve kapalı alanlar ile kullanıldığında GYR verileri FA'yı tahmin etmede yetersiz kalmaktadır. Bu noktada, ACC, CLI ve MAG verilerinin başarısı açıktır. Benzer şekilde, farklı fiziksel aktivitelerdeki manyetik alanlar da değişmektedir. Bu bağlamda, FA'nın lokasyon ile tanınması için manyetometre verilerinin kullanılması bize başarılı sonuçlar vermektedir.

Çok kipli analizlere göre en yüksek doğrulukları aldığımız RF ve DNN algoritmaları ile güncel derin öğrenme algoritmalarını karşılaştırdığımızda Çizelge 8.4'deki sonuçları görmekteyiz. Sonuçları incelediğimizde 7.4 bölümünde önerdiğimiz hibrid CNN+LSTM algoritmasının bütün senaryolarda en yüksek doğruluğu verdiğini görmekteyiz.



Çizelge 8.4 Açık ya da kapalı alan içeren FA için önerilen öğrenme algoritmaları ile sonuçlar

	WSwSA					BSwSA					BSwDA				
	DNN	CNN+LSTM	RNN	CNN	RF	DNN	CNN+LSTM	RNN	CNN	RF	DNN	CNN+LSTM	RNN	CNN	RF
A1	100,00	100,00	99,56	97,49	99,99	79,73	80,31	77,37	75,42	78,03	86,52	90,95	83,88	68,39	88,76
A2	100,00	100,00	99,45	96,49	99,99	65,97	66,56	67,57	61,00	66,90	70,12	82,08	75,23	60,00	78,68
A3	100,00	100,00	99,02	94,82	100,00	73,15	76,64	69,36	63,60	69,39	88,55	85,79	82,77	81,62	83,11
A4	100,00	100,00	99,54	96,58	99,99	60,01	66,32	67,64	63,09	52,51	83,01	77,47	66,53	59,60	67,80
B1	100,00	100,00	99,91	98,78	100,00	72,46	74,30	69,88	56,35	83,29	76,97	85,50	76,46	69,15	84,87
B2	100,00	100,00	99,17	95,89	100,00	57,83	66,63	59,05	56,47	64,97	65,38	64,94	60,54	77,83	68,34
B3	100,00	100,00	99,65	97,25	100,00	75,41	68,77	65,13	61,93	75,19	81,21	88,82	79,19	62,61	86,65
B4	100,00	100,00	99,60	96,79	100,00	80,29	88,03	71,90	60,05	83,74	74,94	83,35	78,75	74,11	87,07
C1	100,00	100,00	99,45	97,33	99,99	73,15	79,45	71,41	56,33	53,63	86,03	81,63	80,63	73,90	69,56
C2	100,00	100,00	98,78	92,44	100,00	56,99	65,02	59,26	58,92	53,61	65,71	65,27	63,68	67,97	67,95
C3	100,00	100,00	99,52	97,90	100,00	78,36	73,53	61,39	56,79	66,81	84,00	78,62	76,37	59,52	83,49
C4	100,00	100,00	99,97	99,87	100,00	77,53	78,15	78,03	61,05	88,29	86,03	80,24	78,73	69,02	80,96
Ort.	100,00	100,00	99,46	96,74	100,00	70,10	73,04	67,33	59,60	68,94	78,36	79,43	74,44	68,67	78,04

Yaş gruplarına göre ortalama doğruluk değerlerine baktığımızda, en yüksek değer A Grubu'nda ortaya çıktı. Grup-C'deki verilerin çok kipli ortalama doğruluğu %74,47, Grup-B'deki verilerin çok kipli ortalama doğruluğu %77,06 ve Grup-A'daki verilerin çok kipli ortalama doğruluğu %78,26'dır. Bu doğruluk değerlerine göre yaş azaldıkça aktivitelerin daha belirgin hale geldiği görülmüştür.

### 8.1.2. Sigma tabanlı öznelik ile sınıflandırma bulguları

Öznelikleri sensör verilerinden çıkardıktan sonra, önerilen yöntemimiz SVM ve RF algoritmaları ile karşılaştırılmıştır. Sigma tabanlı öznelikleri daha fazla değerlendirmek için önerilen DNN modeli sigma tabanlı özneliklerle ve sigma tabanlı öznelikleri kullanmadan eğitilmiştir. Sonuçlar Çizelge 8.5'de sunulmuştur.

5.5.1.2'deki senaryoları kullandığımızda; sonuçlara dayanarak, sigma tabanlı özelliklere sahip olan önerilen DNN modeli, sırasıyla %100, %74.05 ve %81.60 doğruluk ile WSwSA, BSwSA ve BSwDA senaryoları için en iyi sonuçları verir. Sonuçlar, sınıflandırma doğruluğunun BSwSA ve BSwDA senaryoları için WSwSA'ya kıyasla önemli ölçüde azaldığını açıkça göstermektedir. Bunun nedeni, eğitim verilerindeki deneklerin çeşitliliği arttıkça doğruluğun azalmasıdır.

Çizelge 8.5 Farklı yaş ve kişiye göre önerilen yöntemin açık ya da kapalı alan içeren FA için doğruluklarının karşılaştırılması

	WSwSA				BSwSA				BSwDA			
	SVM (%)	RF (%)	Önerilen DNN (%)	Sigma Tabanlı Özellikleri Kullanarak Önerilen DNN (%)	SVM (%)	RF (%)	Önerilen DNN (%)	Sigma Tabanlı Özellikleri Kullanarak Önerilen DNN (%)	SVM (%)	RF (%)	Önerilen DNN (%)	Sigma Tabanlı Özellikleri Kullanarak Önerilen DNN (%)
<b>A1</b>	100,00	99,99	100,00	<b>100,00</b>	66,89	78,03	79,73	<b>81,41</b>	67,17	88,76	86,52	<b>91,86</b>
<b>A2</b>	100,00	99,99	100,00	<b>100,00</b>	42,09	66,9	65,97	<b>67,42</b>	60,75	78,68	70,12	<b>79,15</b>
<b>A3</b>	100,00	100,00	100,00	<b>100,00</b>	48,3	69,39	73,15	<b>77,32</b>	65,02	83,11	88,55	<b>88,56</b>
<b>A4</b>	100,00	99,99	100,00	<b>100,00</b>	43,02	52,51	60,01	<b>63,11</b>	63,29	67,80	83,01	<b>83,41</b>
<b>B1</b>	100,00	100,00	100,00	<b>100,00</b>	57,49	83,29	72,46	<b>75,11</b>	53,93	84,87	76,97	<b>80,44</b>
<b>B2</b>	100,00	100,00	100,00	<b>100,00</b>	40,49	64,97	57,83	<b>62,26</b>	46,27	68,34	65,38	<b>65,55</b>
<b>B3</b>	100,00	100,00	100,00	<b>100,00</b>	46,48	75,19	75,41	<b>75,46</b>	56,45	86,65	81,21	<b>85,71</b>
<b>B4</b>	100,00	100,00	100,00	<b>100,00</b>	56,06	83,74	80,29	<b>87,17</b>	57,89	87,07	74,94	<b>80,43</b>
<b>C1</b>	100,00	99,99	100,00	<b>100,00</b>	43,41	53,63	73,15	<b>80,32</b>	46,30	69,56	86,03	<b>86,52</b>
<b>C2</b>	100,00	100,00	100,00	<b>100,00</b>	33,60	53,61	56,99	<b>60,65</b>	42,25	67,95	65,71	<b>65,91</b>
<b>C3</b>	100,00	100,00	100,00	<b>100,00</b>	45,33	66,81	78,36	<b>79,36</b>	49,10	83,49	84,00	<b>84,51</b>
<b>C4</b>	100,00	100,00	100,00	<b>100,00</b>	50,86	88,29	77,53	<b>78,99</b>	63,78	80,96	86,03	<b>87,10</b>

Yaş gruplarına göre ortalama doğruluk değerleri incelendiğinde, en yüksek değer sigma tabanlı özellikleri kullanan DNN modelimiz ile Grup-A için elde edilir. A Grubu, B Grubu ve C Grubu verilerinin ortalama doğruluğu sırasıyla %86.02, %84.34 ve %85.28'dir. Sonuçlara göre, kişinin yaşı azaldıkça etkinliklerin daha belirgin hale geldiğini gösterebiliriz. Ayrıca, yöntemimiz seçilen diğer yaş gruplarından insanlar için iyi sonuç veriyor. Çünkü, BSwDA'da (Farklı Yaş Grubundaki Kişiler Arasında), eğitim ve test prosedürlerinde farklı yaş grubuna sahip farklı kişilerin örnekleri kullanılmıştır. Farklı yaş gruplarını kullanmanın amacı, yaşlar arasındaki FA farklılıklarını ortadan kaldırmaktır. Çalışmamızdaki Çizelge 8.6'ya göre, WSwSA'nın eğitim ve test prosedürlerini kullandığımızda doğruluk değerleri %100 idi. Ayrıca, BSwSA'nın eğitim ve test prosedürlerini kullandığımızda, A Grubu, B Grubu ve B Grubundaki verilerin ortalama doğruluğu sırasıyla %72.32, %75.00 ve %74.83'tür. Bu sonuçlara göre, tüm yaş gruplarının eğitime katıldığı BSwDA senaryosunda, sadece yaş gruplarının kullanıldığı BSwSA grubundan daha yüksek doğruluk elde edilmiştir. Bu değerler bize farklı yaş gruplarının kullanılmasının doğruluğu

arttırdığını göstermektedir. Ek olarak, sadece gençlerde değil, tüm yaş gruplarının doğruluğunun arttığını göstermektedir.

Çizelge 8.6'da, değerlendirmeler için doğruluk ölçütüne ek olarak kesinlik, duyarlılık ve F1 ölçüleri de dahil edilmiştir. Sigma tabanlı sınıflandırmaya sahip DNN için kesinlik, duyarlılık ve F1 değerleri % 70'in üzerindedir. BSwDA senaryosunu kullanırsak, ortalama kesinlik, duyarlılık ve F1 değerleri düşer. Bu değerler bize doğruluk değerleri ile benzer sonuçlar verir.

Çizelge 8.6 Önerilen Sigma tabanlı öznitelikler ile DNN kullanıldığında ortalama precision ve recall değerleri (%)

WSwSA			BSwSA			BSwDA		
Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1
100,00	100,00	100,00	75,85	82,49	79,03	82,32	85,64	84,00

7.4 bölümünde anlatılan en yüksek doğruluğu aldığımız Hibrid CNN+LSTM algoritmasını, Sigma tabanlı olarak en yüksek doğruluğu aldığımız DNN algoritması ile karşılaştırdığımızda Çizelge 8.7'deki sonuçları görmekteyiz. Tablodaki sonuçlara baktığımızda CNN+LSTM algoritması sigma tabanlı öznitelikler ile kullanıldığında ortalama olarak en yüksek sonucu verdiği görülmektedir.

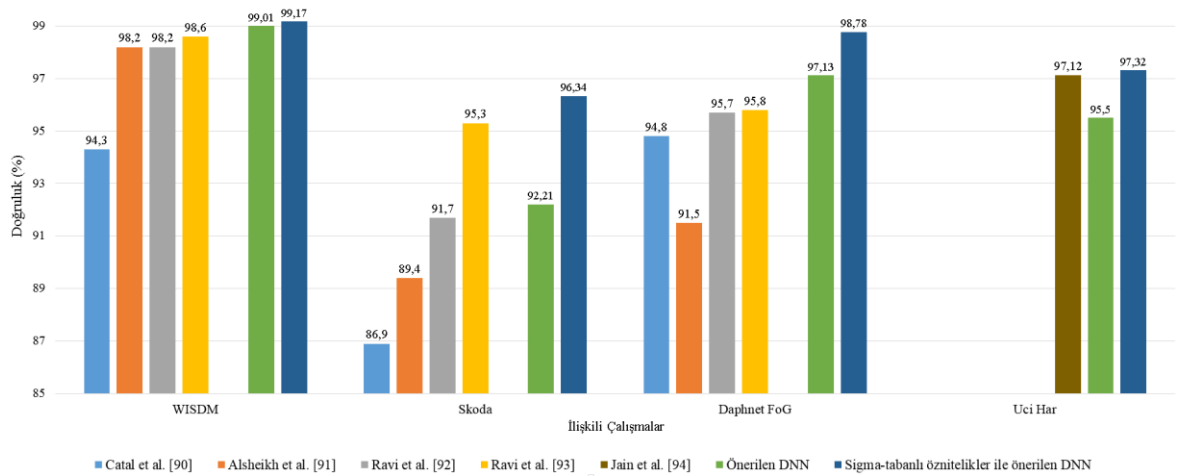
Bu bölümde, Sigma Tabanlı Öznitelik oluşturma ve Üstel hareketli ortalama tabanlı Öznitelik Oluşturmalar ile üretilen yeni özniteliklerin farklı veri kümeleri üzerinde test edilmesi anlatılacaktır. Farklı veri kümelerinde test edilen yöntemler, literatürdeki farklı çalışmalar ile karşılaştırılmıştır. En iyi performans gösteren yöntemlerimizin (Sigma-tabanlı DNN) literatürle karşılaştırılması sağlanmıştır. Bu amaçla, karşılaştırma için WISDM [66], Daphnet Fog [67], Skoda [68] ve UCI HAR [25] performans veri kümeleri kullanılmıştır. Kullanılan bu veri kümeleri ile ilgili ayrıntılı bilgileri Çizelge 4.6'da verilmiştir.

Sigma tabanlı öznitelikleri kullandığımız yöntemimiz ile bu veri kümelerini kullanan beş farklı yöntemle karşılaştırılmıştır. Bu yöntemlerin yaklaşımımızla karşılaştırılması, Şekil 8.2'de gösterilmektedir [99]. Karşılaştırılan yöntemler için tüm veri kümelerinde üstün sonuçlar elde edilmiştir. Skoda ve UCI HAR veri kümelerinde doğruluktaki hafif düşüşün, örneklem sayısının düşük olmasından dolayı olduğu

düşünülmektedir. Ek olarak, bu veri kümesindeki az sayıdaki örnekleme rağmen çok sayıda sınıf vardır. Bunların sonuçları da etkilediği düşünülmektedir. Diğer iki veri kümesinde daha yüksek doğruluk elde edilmiştir. Bununla birlikte, sigma tabanlı özniteliklere sahip olan DNN ile tüm veriler üzerinde en iyi sonuçlar elde edilmiştir. Bunun nedeni, yeni oluşturduğumuz sigma tabanlı özniteliğimizin sensör verilerinin uç (outlier) değerlerini temizlemesidir ve doğruluğu artırmak için bu önemlidir.

Çizelge 8.7 Açık ya da kapalı alan içeren FA sınıflandırma için sigma tabanlı öznitelikleri (özn.) kullanarak DNN ve CNN+LSTM karşılaştırması

	WSwSA		BSwSA		BSwDA	
	Sigma Tabanlı Özn. Kullanarak Önerilen DNN (%)	Sigma Tabanlı Özn. Kullanarak Önerilen CNN+LSTM (%)	Sigma Tabanlı Özn. Kullanarak Önerilen DNN (%)	Sigma Tabanlı Özn. Kullanarak Önerilen CNN+LSTM (%)	Sigma Tabanlı Özn. Kullanarak Önerilen DNN (%)	Sigma Tabanlı Özn. Kullanarak Önerilen CNN+LSTM (%)
A1	100,00	100,00	81,41	81,81	91,86	90,96
A2	100,00	100,00	67,42	66,97	79,15	82,70
A3	100,00	100,00	77,32	77,40	88,56	86,73
A4	100,00	100,00	63,11	66,81	83,41	83,06
B1	100,00	100,00	75,11	75,71	80,44	86,56
B2	100,00	100,00	62,26	68,08	65,55	65,43
B3	100,00	100,00	75,46	69,98	85,71	89,62
B4	100,00	100,00	87,17	88,75	80,43	83,62
C1	100,00	100,00	80,32	80,20	86,52	81,99
C2	100,00	100,00	60,65	65,47	65,91	66,45
C3	100,00	100,00	79,36	74,05	84,51	84,40
C4	100,00	100,00	78,99	79,26	87,10	80,35
Ort	100,00	100,00	74,05	74,54	81,60	81,82



Şekil 8.2 Sigma tabanlı öznitelik kullanarak veri kümelerinin karşılaştırılması

### 8.1.3. EMA öznitelikler kullanıldığında farklı veri kümelerindeki bulguları

Üstel hareketli ortalama tabanlı öznitelikleri kullandığımız yöntemimiz ile bu veri kümelerini kullanan beş farklı yöntemle karşılaştırılmıştır. Bu karşılaştırma esnasında 8.1.2'de anlatılan en yüksek doğruluğu veren yöntem ile de karşılaştırılmıştır. Bu karşılaştırma esnasında ilk önce 6.2.1'de anlatılan EMA ile öznitelikleri oluşturduktan sonra karşılaştırma yapılmıştır. İstenen sonuçlar alınmadığı için 6.2.2'de tanıtılan DEMA yöntemine geçilmiştir. Bu yöntemle öznitelikler oluşturulduktan sonra yeniden testler yapılmıştır. Bu yöntem nispeten tatmin edici sonuçlar verse de yine de yeterli kalmamıştır. Son olarak da 6.2.3 bölümünde anlatılan TEMA yöntemine geçilmiştir. TEMA yöntemi ile öznitelikler üretildikten sonra 8.1.1 bölümünde en başarılı sonuçları aldığımız CNN+LSTM algoritması ile testlerimizi gerçekleştirilmiştir. Çizelge 8.8'de Üstel hareketli ortalamalar ile kullandığımız CNN+LSTM algoritmasının kullanılmasının, literatürdeki diğer veri kümeleri ile karşılaştırması bulunmaktadır. Sensör verisindeki trendleri belirlemek ve sensörlerden gelen hareketin yönünü analiz etmek için kullanılan üstel hareketli ortalamalar ile sensör verisindeki eğilim daha doğru tahmin edilmiştir. Çizelge 8.8'de veri kümelerinin karşılaştırılması yanı sıra TEMA ile EMA ve DEMA karşılaştırması da görülmektedir. TEMA, DEMA ve EMA'ya göre sensör hareketlerine daha duyarlıdır. WISDM ve UCI HAR veri kümelerinde DEMA tabanlı özniteliklerimiz yeterli olmaktadır. Bu iki veri kümesinde DEMA tabanlı öznitelikler ile yüksek sonuç vermesinin sebebi olarak; veri kümelerinin zamansal olarak trende daha uygun olmasından kaynaklanabileceği düşünülmektedir. Diğer veri kümeleri için de hareketin doğru trendini yakalayabilmek için TEMA ile öznitelik çıkarımı yapılmıştır.

Çizelge 8.8 Açık ya da kapalı alan içeren FA sınıflandırması için üstel hareketli ortalama tabanlı öznitelikleri kullanarak veri kümelerinin karşılaştırılması

	WISDM	Skoda	Daphnet FoG	UciHar
Catal vd. [90]	94,3	86,9	94,8	
Alsheikh vd. [91]	98,2	89,4	91,5	
Ravi vd. [92]	98,2	91,7	95,7	
Ravi vd. [93]	98,6	95,3	95,8	
Jain vd. [94]				97,12
Sigma-tabanlı öznitelikler ile önerilen DNN	99,17	96,34	98,78	97,32
EMA tabanlı öznitelikler ile önerilen CNN+LSTM	98,79	93,19	95,60	96,63
DEMA tabanlı öznitelikler ile önerilen CNN+LSTM	<b>99,19</b>	95,24	96,21	<b>97,54</b>
TEMA tabanlı öznitelikler ile önerilen CNN+LSTM	<b>99,31</b>	<b>97,64</b>	<b>99,28</b>	<b>97,86</b>

## 8.2. Physionet Veri Kümesi Bulguları

İlgili veri kümesi çalışmalarında, uyku apnesi bulma işlemi için SpO<sub>2</sub> ve EKG sensörlerinden gelen verilerden ilk olarak Sigma-tabanlı öznitelik oluşturulmuştur. Bunun devamında da 4 değişik öznitelik seçme yöntemi ile öznitelikler oluşturulmuştur. Devamında NB, kNN ve SVM, makine öğrenme methodları olarak tasarlanmıştır. Bu çalışma kapsamında veri kümesi olarak 5.4.3. bölümünde ayrıntılı olarak açıklanan veri kümesi kullanılmıştır.

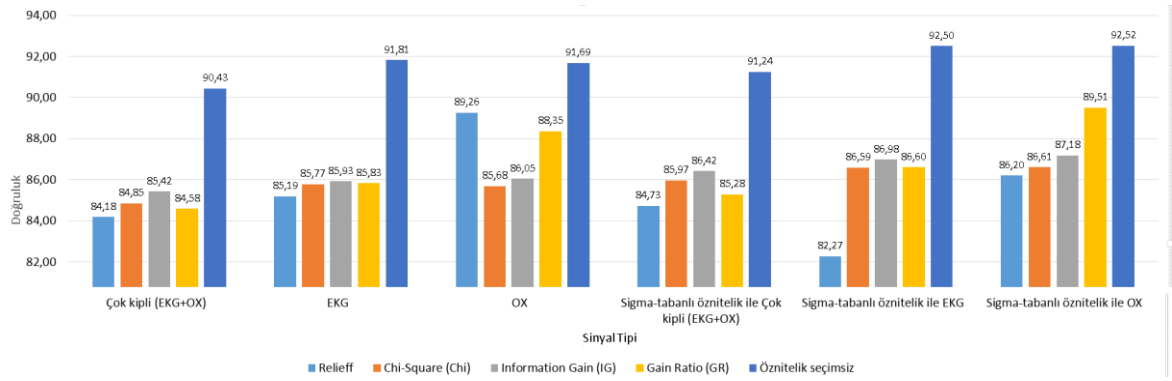
Kullandığımız veri kümesinde EKG ve SpO<sub>2</sub> sensörü için Sigma-tabanlı yeni öznitelikler oluşturulduktan sonra öznitelik seçme işlemi uygulanmıştır. Öznitelik seçme yöntemi aşamasında sınıflandırıcı etiketlerine mesafelerine göre 50 ile 2500 öznitelik içerisinde öznitelik seçme methodu ile analizler yapılmıştır. Örn. İlk 500 yüksek öncelikli öznitelik bulmak amacıyla ilgili seçme methodundan bulunan ağırlıklara göre en yüksek 500'ü kullanılmıştır. Öznitelik seçme yöntemleri olarak Relief, Bilgi Kazancı, Kazanım Oranı ve Ki-Kare yöntemleri kullanılmıştır.

OUA tanıma için NB, kNN ve DVM sınıflandırma yöntemleri kullanılmıştır ve birbirleri içerisinde doğruluk karşılaştırması yapılmıştır. RBF çekirdek fonksiyonu DVM algoritması için kullanılmıştır. Kullanılan çekirdek fonksiyonunun parametrelerini iyileştirmek için de Libsvm Kütüphanesi ile gridsearch methodu kullanılmıştır[11]. Naive Bayes için dağılım olarak Gaussian seçilmiştir. K En Yakın Komşu Algoritması için k değeri 3 seçilmiştir. Bu algoritmadaki uzaklık hesaplama için Öklid yöntemi seçilmiştir. Test aşamasında 10-fold çapraz geçerlilik (k-fold cross validation) methodu uygulanmıştır.

### 8.2.1. Sigma tabanlı öznitelik ile sınıflandırma bulguları

EKG ve SpO<sub>2</sub> sensör verileri ile öncelikle füzyon methodu uygulanmıştır. Daha sonra füzyon edilmiş sensör verileri için ortalama ile standart sapma değerlerine bakarak çalışmamızda ayrıntılı olarak anlatılan, yeni veriler oluşturulmuştur. Bu veriler ile oluşturulan öznitelikler oluşturulduktan sonra öznitelik seçme methodları uygulanmıştır. Daha sonra da makine öğrenme algoritması işlemi başlanmıştır. Bu çalışmada ayrıntılı olarak anlattığımız sigma-tabanlı öznitelik oluşturma yönteminin başarımını bulmak amacıyla uyku apnesi şiddetine göre kişiler (yani yüksek apne, düşük apne ya da apnesiz) birbirinden ayrılmıştır.

Şekil 8.3'de bu tez çalışmasında önerilen sigma-tabanlı öznitelikler ile farklı öznitelik seçme methodlarının başarımlarının karşılaştırması görülmektedir [100]. İlgili grafikte öznitelik seçimi uygulanmadığında, makine öğrenme doğruluğu, kullanılan farklı sensörler için de en yüksek olduğu görülmektedir. Yine benzer şekilde sigma-tabanlı öznitelikleri kullanıldığında başarımın tek ve çok kipli çalışmalarda arttığı görülmüştür. Öznitelik seçme methodları analizlerin hepsinde çok benzer doğruluklar vermiştir; bununla beraber öznitelik seçimi yapılmayan methoddan başarımı daha düşüktür. Bu sebeplerden dolayı, veri azlığı ile sağladığı işlem yapma hızı düşüklüğü sebebi ile öznitelik seçme yöntemi tercih edilebilir. Çok kipli analizlerde Bilgi kazancı yöntemi diğer methodlara göre daha yüksek başarımlar vermiştir.



Şekil 8.3 Uyku apnesi sınıflandırması için sensor verisi, öznitelik seçme yöntemi ile doğruluk karşılaştırması

## 9. TARTIŞMA VE SONUÇ

Bu çalışmada, sağlık bilişimi alanında, yaşlı ve bakıma muhtaç kişilerin, sağlık durumlarının ve günlük aktivitelerinin takibi amacıyla, farklı sensörlerden toplanan verilerden anlamlı bilgilerin elde edilebilmesi için, çok kipli ve füzyon tabanlı analizlere dayalı akıllı yöntemler geliştirilmiştir. Giyilebilir sensörler olarak nabız, oksijen doygunluğu kullanılmıştır. Bu sensörlerden toplanan veriler ile aritmi, uyku apnesi hastalıklarının takibi ve erken tanınmasının yanı sıra, cep telefonu içerisinde bulunan sensörlerden ACC, CLI, MAG ve GYR sensörleri ile izlenen kişinin fiziksel aktivitesi incelenmiştir.

İlk olarak, açık ve kapalı alan konumu ile FA'nın tanınması için yeni bir yöntem önerilmiştir. Bu amaçla, on altı saatlik, sensörlerden (ACC, GYR, MAG ve CLI) ayrıca FA ve ortamı ile etiketlenmiş oksimetre verilerinden oluşan yeni oluşturulan bir veri kümesi tanıtılmıştır. Sensör verilerinin aykırı değerlerinin etkisini azaltmak için, sigma tabanlı özellikler tanıtıldı ve aykırı verileri veriden çıkarmak için sigma tabanlı bir özellik seçimi algoritması önerilmiştir. Önerilen DNN modeli, tanıtılan özellikleri değerlendirmek için kullanılmıştır. Önceki çalışmalarla karşılaştırıldığında, önerilen sigma tabanlı DNN tüm veri kümelerinde daha iyi doğruluk sağlar.

Bu çalışmada, sensör sinyalleri açısından birbirinden önemli ölçüde farklı olan yürüme, koşma, ayakta durma ve oturma olmak üzere dört temel aktivite ele alınmıştır. FA kapalı alanda yapılırken, kapalı alan olarak klimalı olmayan bir alan kullanılmıştır. Benzer şekilde, FA kapalı alanda yapıldığında, seyrek ağaçlı ve koşu parkurlu ortamlar seçilmiştir. Gelecekteki çalışmalarda, farklı çevre koşullarında yapılabilecek daha ayrıntılı faaliyetler incelenebilir.

Sunulan sonuçlar, FA'nın öngörülmesinde ağırlıklı olarak yerin zorluklarını karakterize etmeye hizmet edebilir. Birden fazla sensör kullanmak veri kümesindeki aykırı değerleri artırabilir ve aykırı kümeleri tanımlamak ve tanımlamak için her veri kümesinin özellikleri nedeniyle kesin bir yol yoktur. Bu aykırı değerler, makine öğrenimi eğitim sürecinin başarısını etkiler. Bu nedenle, veri kümesindeki aykırı değerlerin kaldırılması, makine öğrenme algoritmalarının başarısını artırabilir.

Bu tez çalışması sonrasında planlanan çalışmalar, farklı füzyon stratejilerinin ve öznelik seçim yöntemlerinin araştırılmasında yatmaktadır. Her sensör, sensör belirli algoritmaları ve gözlemlere dayalı sınıf hakkındaki bilgileri ortaya çıkarmaya



yardımcı olmaktadır. Bu nedenle, sınıf bilgisini içerisinde daha fazla barındıran sensör verisinin, füzyon esnasında ağırlığının artırılmasının, sınıflandırma başarımına katkısının daha fazla olacağı düşünülmektedir. Bu amaçla, veri füzyonunda amaca yönelik en uygun yöntemin seçilmesi planlanmaktadır.

Uyku apnesi sınıflandırma görevi için çok kipli bir yaklaşım sunulmuştur. Spesifik olarak, elektrokardiyografi (EKG) ve periferik oksijen (SpO<sub>2</sub>) sinyallerinin doygunluğundan öznitelikleri ayıklarız ve bunları eğiticiye vermeden önce uygun bir füzyon yöntemiyle birleştiririz. Bu çalışmada anlatılan yöntemlerimizin etkinliğini göstermek için farklı test senaryoları göz önüne alınarak NB, SVM ve kNN makine öğrenme algoritmaları seçilmiştir. PhysioNet veri kümesi kullanarak yaptığımız çalışmalardaki sonuçlarımız, öznitelik füzyonu ile çok kipli yaklaşımın tüm analizler içerisinde %96,64 ile SVM methodu en iyi maline öğrenme başarımını vermektedir. Tek kipli yaklaşım kullanılırken de, doğruluk %95,88'e düşmektedir. WSwSS ve BSwSS senaryoları için genel iyileştirmelerin küçük olduğunu, ancak sonucun algoritmanın genelleştirme yeteneğini gösterdiği BSwDS senaryosu göz önüne alındığında, iyileştirmenin %3,04'lük bir artışla önemli olduğunu gözlemlenmiştir. Öte yandan, tek kipli yaklaşım, özellik füzyonu tarafından sunulan özellik boyutluluğunu korumak için tercih edilebilir. Ek olarak, SVM makine öğrenme methodu diğer methodlara göre daha yüksek doğruluk vermektedir.

Ayrıca uyku apnesi sınıflandırması için, sınıf bilgisinin elde edileceği özniteliklerin bazılarının kullanımı, sınıflandırma başarısını artırmada başarılı olamamaktadır. Bu amaçla, sınıflandırma işleminden önce farklı öznitelik seçme yöntemleri uygulandığında başarımı daha kısa sürede ve daha yüksek değerde çıkmaktadır. Benzer şekilde, uyku apnesi (OSA) sınıflandırması için sigma-tabanlı bir yaklaşım sunulmaktadır. EKG ve SpO<sub>2</sub> verilerinden sigma-tabanlı öznitelikler oluşturulmuştur. Oluşturulan yeni özniteliklerin başarımını göstermek için farklı test senaryolarını oluşturulmuştur. Daha sonra da NB, DVM ve kNN makine öğrenme algoritmaları bu verilere uygulanmıştır. PhysioNet veri kümesindeki sonuçlarımız, öznitelik füzyon işlemi ile çok kipli methodun, tüm analizlerde %95,52'lik başarımla SVM metodu ile en yüksek başarımı verdiğini göstermektedir. Diğer taraftan, tek kipli yöntem, çok kipli analizde oluşan öznitelik boyutunu probleminden kurtulmak için seçilebilir.

Son olarak, bir kişinin açık ve kapalı alan konumlarının sınıflandırılması için bir topluluk yöntemi sunulmuştur. SpO<sub>2</sub> öznitelikleri çıkarılmıştır ve sınıflandırma için NB, SVM, RF ve kNN algoritmaları kullanılmıştır. Önerilen veri kümemizden elde edilen gerçek klinik örnekler hakkındaki sonuçlarımız, en iyi sınıflandırma doğruluğunun, RF senaryosuyla tüm senaryolar arasında ortalama % 69,34 doğrulukla elde edildiğini göstermektedir. Oksijen doygunluğunu kullanmak, bir insanın iç veya dış mekânlarını bulmak için, özellikle de 51-65 yaş arasındaki insanlar için etkilidir. Diğer yandan, eğer insanın yaşı artarsa, sınıflandırma doğruluğu da artar. Ayrıca, insanın yaşı artarsa, ortalama oksijen doygunluğu azalır. Sınıflayıcılar arasında, RF algoritması diğerlerinden daha üstündür. Gelecekteki çalışmalarımız derin öğrenme yöntemlerini kullanarak daha sağlam sınıflandırma stratejilerinin araştırılmasına dayanmaktadır.

Bu tez kapsamında uygulanan tüm çalışmalar ve yapılan analizlere göre, sigma tabanlı veya üstel hareketli ortalama tabanlı öznitelik çıkarım yöntemleri uygulandıktan sonra sensör verileri sınıflandırılması yüksek başarımla sonuçlandırılmıştır. Yapılan çalışmalarda bulunan başarımın yükseltilmesi amacıyla için, oluşturulan öznitelik oluşturma yönteminin yeni veri kümeleri ve yeni sınıflandırma yöntemleri ile zenginleştirilmesi gerektiği düşünülmektedir.

Bu tez çalışmasında önerilen öznitelik oluşturma yöntemlerinin, gelecekte daha da yaygınlaşacak olan giyilebilir sensörler kullanılarak oluşturulan kişiselleşmiş tıp uygulamaları içerisinde de kullanılabileceği düşünülebilir. Kişiselleşmiş tıp çalışmalarının ortasında olan insanın, sensör verilerinin kullanılarak hesaplanması ile bu yeni yaklaşım dâhilinde daha iyi temsil edebileceği düşünülebilir. Kişiselleşmiş tıp çalışmaları ve deneklerin yakından takip edilmesi için önerilen sensör analizleri ve geliştirilen öznitelik oluşturma yöntemlerinin kullanılabileceği düşünülmektedir.

## KAYNAKLAR

- [1] Chan, M., Estève, D., Fourniols, J.Campo, E., "Smart wearable systems: future challenges", *Artificial intelligence*, 56(3), 137-156. , 2012
- [2] Barnes, K., "Health wearables: Early days", PwC Health Research Institute Report., 2014
- [3] Bonato P: Wearable sensors and systems. From enabling technology to clinical applications. *IEEE Eng Med Biol Mag* 2010, 29:25-36.
- [4] Chan A.M., Nandakumar S., Nima F., and Ravi N. ,Wireless Patch Sensor for Remote Monitoring of Heart Rate, Respiration, Activity, and Falls (35th Annual International Conference of the IEEE EMBS Osaka, Japan, 3 - 7 July, 2013
- [5] Caudill TS, Lofgren R, Jennings CD, Karpf M: Commentary: Health care reform and primary care: training physicians for tomorrow's challenges. *Acad Med* 2011, 86:158-160.
- [6] Bayat A. vd., "A study on human activity recognition using accelerometer data from smartphones," *Proc. Comput. Sci.*, vol. 34, pp. 450–457, Aug. 2014.
- [7] Yeh H. P., Stone J. A. vd., "Physical and Emotional Benefits of Different Exercise Environments Designed for Treadmill Running," *Int J Environ Res Public Health*, vol. 14, no. 7, p. 752, Jul. 2017.
- [8] Zhou P. vd., "IODetector: a generic service for indoor outdoor detection," in *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems, USA*, pp. 361–362.
- [9] Radu V., Katsikouli P., Sarkar R., ve Marina M. K., "A semi-supervised learning approach for robust indoor-outdoor detection with smartphones," in *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems, ACM, USA*, pp. 280–294.
- [10] "Clinical User of Pulse Oximetry," 2010. [Online]. Available: <http://www.copdalert.com/OximetryPG.pdf>. [Accessed: Jan. 13, 2019].
- [11] Kansiz A. O. vd., "Selection of timedomain features for fall detection based on supervised learning," presented at *World Congr. Eng. Comput. Sci.*, 2013.

- [12] Siirtola P. ve Roning J., "Ready-to-use activity recognition for smartphones," IEEE Symp. Comput. Intell. Data Min. (CIDM), Apr. 2013, Singapore.
- [13] Seneviratne S., Hu Y. vd., "A Survey of Wearable Devices and Challenges," IEEE Communications Surveys & Tutorials, vol. 19, no. 4, 2017.
- [14] Su X., Tong H., ve Ji P., "Activity Recognition with Smartphone Sensors," Tsinghua Science and Technology, vol. 19, no. 3, pp. 235–249, 2014.
- [15] Bulling A. vd., "A tutorial on human activity recognition using body-worn inertial sensors," ACM Computing Surveys (CSUR), vol. 46, p. 33, 2014.
- [16] Yang J. B. vd., "Deep convolutional neural networks on multichannel time series for human activity recognition," in IJCAI, Argentina, 2015, pp. 25–31.
- [17] Yang Q., "Activity recognition: Linking low-level sensors to high-level intelligence," in IJCAI, 2009, pp. 20–25.
- [18] He J., Zhang Q., Wang L. ve Pei L., "Weakly Supervised Human Activity Recognition From Wearable Sensors by Recurrent Attention Learning," in IEEE Sensors Journal, vol. 19, no. 6, pp. 2287-2297, 15 March 2019.
- [19] Ordóñez F. J. ve Roggen D., "Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition," Sensors, vol. 16, no. 1, p. 115, 2016.
- [20] Ha S. ve Choi S., "Convolutional neural networks for human activity recognition using multiple accelerometer and gyroscope sensors," in Neural Networks (IJCNN), 2016 International Joint Conference on. IEEE.
- [21] Radu V., Katsikouli P., Sarkar R., ve Marina M. K., "A Semi-supervised Learning Approach for Robust Indoor-outdoor Detection with Smartphones," in Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems. ACM, USA, 2014, pp. 280–294.
- [22] Wang Z., Wu D., Chen J., Ghoneim A., ve Hossain M. A., "A Triaxial Accelerometer-Based Human Activity Recognition via EEMD-Based Features and Game-Theory-Based Feature Selection," IEEE Sensors Journal, vol. 16, no. 9, pp. 3198–3207, 1 May, 2016.

- [23] Lv M. vd., "Bi-View Semi-Supervised Learning Based Semantic Human Activity Recognition Using Accelerometers," *IEEE Transactions on Mobile Computing*, vol. 17, no. 9, pp. 1991–2001, 1 Sept. 2018.
- [24] Fullerton E., Heller B., ve Munoz-Organero M., "Recognizing Human Activity in Free-Living Using Multiple Body-Worn Accelerometers," *IEEE Sensors Journal*, vol. 17, no. 16, pp. 5290–5297, 2017.
- [25] Jain A. ve Kanhangad V., "Human Activity Classification in Smartphones Using Accelerometer and Gyroscope Sensors," *IEEE Sensors Journal*, vol. 18, no. 3, pp. 1169–1177, 1 Feb. 2018.
- [26] Incel, O. Durmaz, M. Kose, ve C. Ersoy, "A review and taxonomy of activity recognition on mobile phones," *BioNanoScience*, vol. 3.2, pp. 145–171, 2013.
- [27] Esfahani M. I. M. ve Nussbaum M. A., "A "Smart" undershirt for tracking upper body motions: task classification and angle estimation," *IEEE Sensors Journal*, vol. 18, no. 18, pp. 7650–7658, 2018.
- [28] Mukhopadhyay S. C., "Wearable sensors for human activity monitoring: A review," *IEEE Sensors Journal*, vol. 15.3, pp. 1321–1330, 2015.
- [29] Shoaib, M. vd. "A survey of online activity recognition using mobile phones," *Sensors*, vol. 15.1, pp. 2059–2085, 2015.
- [30] Schrack J. A. vd., "Assessing daily physical activity in older adults: unraveling the complexity of monitors, measures, and methods," *J. Gerontol. A Biol. Sci. Med. Sci.*, vol. 71.8, pp. 1039–1048, 2016.
- [31] American Academy of Sleep Medicine (AASM) Task Force, "Sleep-related breathing disorders in adults: Recommendations for syndrome definition and measurement techniques in clinical research," *Sleep*, vol. 22, pp. 667–689, 1999. J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp. 68–73.
- [32] Insufficient Sleep Is a Public Health Epidemic. [cited February 2, 2013]; Available from: <http://www.cdc.gov/features/dssleep/>.

- [33] Penzel, T., vd., Systematic comparison of different algorithms for apnoea detection based on electrocardiogram recordings. *Medical and Biological Engineering and Computing*, 2002. 40(4): p. 402-407.
- [34] Lavie, L., Obstructive sleep apnoea syndrome  $\pm$  an oxidative stress disorder. *Sleep Medicine Reviews*, 2003. 7(1): p. 35-51.
- [35] Mendez M. O., vd., Automatic screening of obstructive sleep apnea from the EKG based on empirical mode decomposition and wavelet analysis. *Physiol. Meas.*, 2010. 31: p. 273-289.
- [36] Kakkar R. K. ve Berry R. B., "Positive airway pressure treatment for obstructive sleep Apnea," *Chest*, vol. 132, pp. 1057–1072, 2007.K. Elissa, "Title of paper if known," unpublished.
- [37] Armon C, Johnson G.K, Roy A, Nowack W.J, "Polysomnograph". Available: <http://emedicine.medscape.com/article/1188764-overview>
- [38] Chazal P, Penzel T, and Heneghan C, "Automated detection of obstructive sleep apnoea at different time scales using the electrocardiogram". *Physiological Measurement*. July 2004; 25: 967-983.
- [39] Penzel T., McNames J., de Chazal P., Raymond B., Murray A., Moody G., "Systematic comparison of different algorithms for apnoea detection based on electrocardiogram recordings", *Medical and Biological Engineering and Computing*. 2002; 40:402–407.
- [40] Yilmaz B, Asyali MH, Arikan E, Yetkin S, Ozgen F. "Sleep stage and obstructive apneic epoch classification using single-lead ECG". *Biomed Eng Online*. 2010 Aug 19;9(1):39.
- [41] Khandoker A. H. vd., "Support vector machines for automated recognition of obstructive sleep apnea syndrome from EKG recordings," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 1, pp. 37–48, Jan. 2009.
- [42] Shouldice R. vd., "Detection of obstructive sleep apnea in pediatric subjects using surface lead electrocardiogram features." *Sleep*, vol. 27, no. 4, pp. 784–792, Jun. 2004.

- [43] Mendez M. O. vd., "Sleep apnea screening by autoregressive models from a single EKG lead," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 12, pp. 2838–2850, Dec. 2009.
- [44] Bsoul M. vd., "Apnea medassist: Real-time sleep apnea monitor using single-lead EKG," *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, no. 3, pp. 416–27, May 2011.
- [45] Isa S. M. vd., "Sleep apnea detection from EKG signal: Analysis on optimal features, principal components, and nonlinearity," in *Proc. 5th Int. Conf. Bioinformat. Biomed. Eng.*, Wuhan, China, 2011, pp. 1–4.
- [46] Oliver N. ve Flores-Mangas F., "HealthGear: A real-time wearable system for monitoring and analyzing physiological signals," in *Proc. IEEE Int. Workshop Wearable Implantable Body Sens. Netw. (BSN)*, Apr. 2006, p. 4.
- [47] Fu-Chung Y., Behbehani K., Lucas E., Burk J., ve Axe J., "A noninvasive technique for detecting obstructive and central sleep Apnea," *IEEE Trans. Biomed. Eng.*
- [48] Khandoker A. H., Palaniswami M., ve Karmakar C., "Support vector machines for automated recognition of obstructive sleep Apnea syndrome from EKG recordings," *IEEE Trans. Inf. Technol. Biomed.*, vol 13, no. 1, pp. 37–48, Jan. 2009.
- [49] Patangay A., Vemuri P., ve Tewfik A., "Monitoring of obstructive sleep Apnea in heart failure patients," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBS) Cit□ 1/2 Int.*, Lyon, France, Aug. 2007, pp. 1043–1046.
- [50] Raymond B., Cayton R. M., Bates R. A., ve Chappell M. J., "Screening for obstructive sleep Apnea based on the electrocardiogram—The computers in cardiology challenge," *Comput. Cardiol.* , vol. 27, pp. 267–270, 2000.
- [51] de Chazal P., Heneghan C., Sheridan E., Reilly R., Nolan P., ve O'Malley M., "Automated processing of the single-lead electrocardiogram for the detection of obstructive sleep Apnea," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 6, pp. 686–696, Jun. 2000
- [52] zephyranywhere.com, 'Application Notes & White Papers', 2016 [Online]. Available: <https://www.zephyranywhere.com/zephyrlabs/white-papers> [Accessed: 29-Oct-2016]

- [53] Sani A.S., Islam A.K.M., Mahrin N., Al-Mamoon I., Baharun S., Komaki S., Biomedical Engineering and Sciences (IECBES), 2014 IEEE Conference on, 159-164
- [54] Ravish D. K., Shanthi K. J., Shenoy N.R., Nisargh S., Heart function monitoring, prediction and prevention of Heart Attacks: Using Artificial Neural Networks, Proceedings of 2014 International Conference on Contemporary Computing and Informatics, IC3I 2014 01/2015; DOI: 10.1109/IC3I.2014.7019580
- [55] Mokhlespour E.M., ve Maury N., "Preferred placement and usability of a smart textile system vs. inertial measurement units for activity monitoring," Sensors, vol. 18.8, p. 2501, 2018.
- [56] "Innovo Oximeter," [Online]. Available: <https://innovomedical.com/collections/prescription-medical-devices/products/m50f>
- [57] "Samsung Galaxy Note 3," [Online]. Available: [www.samsung.com/uk/smartphones/galaxy-note-3-n9005/SM-N9005ZKEBTU/](http://www.samsung.com/uk/smartphones/galaxy-note-3-n9005/SM-N9005ZKEBTU/)
- [58] "Samsung Galaxy S4," [Online]. Available: [www.samsung.com/uk/smartphones/galaxy-s4-i9505/GT-I9505ZKABTU/](http://www.samsung.com/uk/smartphones/galaxy-s4-i9505/GT-I9505ZKABTU/)
- [59] Bathilde J. B. vd., "Continuous heart rate monitoring system as an IoT edge device," Sensors Applications Symposium (SAS) 2018 IEEE, pp. 1–6, 2018.
- [60] Cleland I. vd., "Optimal placement of accelerometers for the detection of everyday activities," Sensors, vol. 13.7, pp. 9183-9200, 2013.
- [61] Boerema S. T. vd., "Optimal sensor placement for measuring physical activity with a 3D accelerometer," Sensors, vol. 14.2, pp. 3188–3206, 2014.
- [62] Özdemir A. T., "An analysis on sensor locations of the human body for wearable fall detection devices: Principles and practice," Sensors, vol. 16.8, p. 1161, 2016.
- [63] Amerini I., Becarelli R., Caldelli R., Melani A., Niccolai M., "Smartphone Fingerprinting Combining Features of On-Board Sensors", IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 12, NO. 10, OCTOBER 2017



- [64] Li P. vd., "An Automatic User-Adapted Physical Activity Classification Method Using Smartphones," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 3, pp. 706–714, Mar. 2017.
- [65] Otter D. vd., "Speed related changes in muscle activity from normal to very slow walking speeds," *Gait & posture*, vol. 19, pp. 270–8, 2004.
- [66] "How to Implement a Neural Network," [Online]. Available: [peterroelants.github.io/posts/neural\\_network\\_implementation\\_intermezzo02](https://peterroelants.github.io/posts/neural_network_implementation_intermezzo02)[Accessed: Feb. 07, 2018].
- [67] Bachlin M. vd., "Wearable assistant for Parkinson's disease patients with the freezing of gait symptom," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 2, pp. 436–446, Mar. 2010.
- [68] Zappi P. vd., "Activity recognition from on-body sensors: Accuracy power trade-off by dynamic sensor selection," in *Wireless Sensor Networks*. Berlin, Germany: Springer, 2008, vol. 4913, ch. 2, pp. 17–33.
- [69] Duda R. O., Peter E. Hart, and Stork D. G., *Pattern classification*. Wiley, 2001.
- [70] Cortes C. and Vapnik V., "Support-Vector Networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
- [71] Chang, C. C., & Lin, C. J. (2011). LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3), 27.
- [72] Breiman L. (2001) "Random Forests" *Machine Learning*, 45, 5-32.
- [73] Nair V., Hinton G. E.. (2010). Rectified linear units improve restricted Boltzmann machines. Presented at *Int. Conf. Mach. Learn.* ,807–814.
- [74] Glorot X. and Bengio Y., (2010). Understanding the difficulty of training deep feedforward neural networks. In *Aistats*. Vol. 9, 249–256.
- [75] Bengio Y., (2009). Learning deep architectures for AI. *Found. Trends® Mach. Learn.* 2(1), 1–127.
- [76] Lecun Y., Bengio Y., Hinton G., Deep learning, *Nature* 521 (7553) (2015) 436–444

- [77] Wang J. vd., Deep learning for sensor-based activity recognition: A Survey, *Pattern Recognition Letters* (2018)
- [78] Hasim S., Andrew S., Françoise B., Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling, *Neural and Evolutionary Computing*, 2014
- [79] Tianyuan L., Jinsong B., Junliang W. and Yiming Z., A Hybrid CNN–LSTM Algorithm for Online Defect Recognition of CO<sub>2</sub> Welding, *Sensors* 2018, 18(12), 4369; doi.org/10.3390/s18124369
- [80] Colby, Robert W., *The Encyclopedia of Technical Market Indicators*, Second Edition, McGraw-Hill Publishing, 2003.
- [81] Kira K., and Rendell L.,(1992). The Feature Selection Problem: Traditional Methods and a New Algorithm. *AAAI-92 Proceedings*.
- [82] Liu H., Setiono R., 1995. Chi2: Feature Selection And Discretization Of Numeric Attributes. In: *Proceedings Of The IEEE 7th International Conference On Tools With Artificial Intelligence* 338-391.
- [83] Alhaj, T. A., Siraj, M. M., Zainal, A., Elshoush, H. T., & Elhaj, F. (2016). Feature Selection Using Information Gain for Improved Structural-Based Alert Correlation. *PloS one*, 11(11), e0166017.
- [84] Memiş G., Sert M.. (2017, Feb). Multimodal Classification of Obstructive Sleep Apnea Using Feature Level Fusion. Presented at 2017 IEEE 11th International Conference on Semantic Computing.
- [85] Bersch S. D. vd., "Sensor data acquisition and processing parameters for human activity classification," *Sensors*, vol. 14.3, pp. 4239–4270, 2014.
- [86] Kwapisz J. R. vd., "Activity recognition using cell phone accelerometers," *ACM SIGKDD Explor. Newslett*, vol. 12, no. 2, pp. 74–82, 2010.
- [87] Banos, O., Garcia, R., Holgado, J. A., Damas, M., Pomares, H., Rojas, I., Saez, A., Villalonga, C. mHealthDroid: a novel framework for agile development of mobile health applications. *Proceedings of the 6th International Work-conference on Ambient Assisted Living an Active Ageing (IWAAL 2014)*, Belfast, Northern Ireland, December 2-5, (2014).

- [88] Chanyaswad T., Chang J.M., Kung S. Y. A compressive multi-kernel method for privacy-preserving machine learning. *Neural Networks (IJCNN)*, International Joint Conference on. 2161-4407. 14-19 May 2017
- [89] Kutlay M. A. and Gagula P. S. Application Of Machine Learning In Healthcare: Analysis On MHEALTH Dataset. *Southeast Europe Journal of Soft Computing* Vol.4 No.2Sep. 2015(46-51)
- [90] Catal C., S. Tufekci, E. Pirmit, and G. Kocabag, "On the use of ensemble of classifiers for accelerometer-based activity recognition," *Appl. Soft Comput.*, vol.37, pp. 1018–1022, Dec. 2015. M. A.
- [91] Alsheikh M.A. vd., "Deep activity recognition models with triaxial accelerometers," Nov. 2015.[Online].Available: <http://arxiv.org/abs/1511.04664>
- [92] Ravi D. vd., "Deep learning for human activity recognition: A resource efficient implementation on low-power devices," in *Proc. 2016 IEEE 13th Int. Conf. Wearable Implantable Body Sensor Netw.*, Jun. 2016, pp. 71–76.
- [93] Ravi D., Wong C., Lo B. vd., "A deep learning approach to onnode sensor data analytics for mobile or wearable devices," *IEEE J Biomed Health Informat.*, vol.21, pp. 56–64, 2017.
- [94] Jain A. and Kanhangad V., "Human Activity Classification in Smartphones Using Accelerometer and Gyroscope Sensors," *IEEE Sensors Journal*, vol. 18, no. 3, pp. 1169–1177, 1 Feb. 2018.
- [95] Goldberger A.L., Amara L.A.N, Glass L., Hausdorff J.M., Ivanov P., Mark R.G., Mietus J.E., Moody G.B., Peng C.K., Stanley H.E. *PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation* 101(23):e215- e220 2000 (June 13).
- [96] Cheng S., Mehrdad N., Gopal G. and Lakshman T., "Apnea MedAssist II: A smart phone based system for sleep apnea assessment", 2013 IEEE International Conference on Bioinformatics and Biomedicine, pp. 572-577, 2013
- [97] Baile X., and Hlaing M., "Real-Time Sleep Apnea Detection by Classifier Combination", *IEEE transactions on information technology in biomedicine*, vol. 16, no. 3, may 2012

- [98] Lang S., Bravo-Marquez F., Beckham C., Hall M., Frank E., "WekaDeeplearning4j: A deep learning package for Weka based on Deeplearning4j", Knowledge-Based Systems, Volume 178, 2019, Pages 48-50
- [99] Memis G. and Sert M., "Detection of Basic Human Physical Activities With Indoor–Outdoor Information Using Sigma-Based Features and Deep Learning," in IEEE Sensors Journal, vol. 19, no. 17, pp. 7565-7574, 1 Sept.1, 2019. doi: 10.1109/JSEN.2019.2916393
- [100] Memis G. and Sert M., (2019), "Classification of Obstructive Sleep Apnea using Multimodal and Sigma-based Feature Representation," IEEE 27th Signal Processing and Communications Applications Conference (SIU 2019), Apr. 24-26, 2019, Sivas, Turkey
- [101] Memis G. and Sert M., "Classification of Indoor-Outdoor Location using Blood Oxygen Saturation Signal," 2018 IEEE 8th International Conference on Consumer Electronics - Berlin (ICCE-Berlin), Berlin, 2018, pp. 1-2.
- [102] Memis G. and Sert M., "The effectiveness of feature selection methods on physical activity recognition," 2018 26th Signal Processing and Communications Applications Conference (SIU), Izmir, 2018, pp. 1-4.
- [103] Memis G. and Sert M. and Yazıcı A., "Leveraging multimodal and feature selection approaches to improve sleep apnea classification performance," 2017 25th Signal Processing and Communications Applications Conference (SIU), Antalya, 2017, pp. 1-4.
- [104] Memis G. and Sert M., "Multimodal Classification of Obstructive Sleep Apnea Using Feature Level Fusion," 2017 IEEE 11th International Conference on Semantic Computing (ICSC), San Diego, CA, 2017, pp. 85-88.